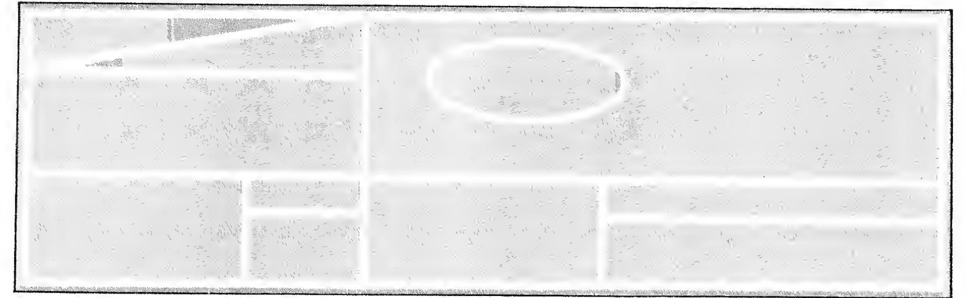


ÁLGEBRA LINEAL



ÁLGEBRA LINEAL

SEGUNDA EDICIÓN



STANLEY I. GROSSMAN

University of Montana

Coordinador de la traducción:
M. en C. Sergio Vargas Galindo

Traductores:
Dr. José Luis Farah
M. en C. Javier Alagón Cano
Mat. Carlos Muñoz Abogado
Mat. Jesús Noriega Rivero
Instituto Tecnológico Autónomo de México (ITAM)

Revisor Técnico:
Ing. Francisco Paniagua Bocanegra
Universidad Nacional Autónoma de México (UNAM)



Grupo Editorial Iberoamérica



Río Atoyac No. 32 - 06500 México, D.F. — Tels. 2113128, 5530798

ÁLGEBRA LINEAL - Segunda edición

Versión en español de la obra *Elementary Linear Algebra* – Third Edition, por Stanley I. Grossman.
Edición original en inglés publicada por Wadsworth, Inc.,

ISBN 0-534-07422-7

ISBN 968-7270-39-X
Impreso en México

Editor: Nicolás Grepe P.
Productor: Oswaldo Ortiz R.
Cubierta: Louis Neiheisel

Grupo Editorial Iberoamérica, S.A. de C.V.
Río Atoyac No. 32 - Col. Cuauhtémoc - 06500 México, D.F.
Apdo. 5-192 - Tels. 2113128, 5530798
Reg. CNIEM 1382

A Kerstin, Erik y Aaron

Prólogo



Hasta hace unos veinte años, sólo estudiaban Álgebra Lineal los estudiantes con graduación en matemáticas y física y quienes, por trabajar en áreas como la estadística matemática de varias variables, necesitaban poseer conocimientos de teoría de matrices. Ahora se estudia Álgebra Lineal en una amplia gama de disciplinas debido a la presencia de las computadoras de alta velocidad y al desarrollo en la aplicación de las matemáticas en áreas tradicionalmente no técnicas.

Al escribir este texto tuve presentes dos objetivos. Traté de hacer que un amplio temario de Álgebra Lineal fuera accesible a una amplia variedad de estudiantes que necesitaban sólo conocimientos básicos de álgebra de bachillerato. Como muchos estudiantes ya habrían cursado un año de cálculo, he incluido también muchos ejemplos de esta materia, los cuales se indican con el símbolo □. Una sección opcional (la Sección 6.7) requiere del cálculo, pero *éste no es requisito indispensable.*

Mi segundo objetivo fue convencer a los estudiantes de la importancia del Álgebra Lineal para sus campos de estudio. De modo que, especialmente en los primeros capítulos, los ejemplos se tomaron de una diversidad de disciplinas. Tales ejemplos son necesariamente breves pero representativos del mundo de las matemáticas. Ejemplos más detallados y extensos pueden encontrarse en la obra complementaria *Aplicaciones de Álgebra Lineal.*

En este texto utilicé un planteamiento gradual. Los Capítulos 1 y 2 contienen material analítico básico común a la mayoría de los textos elementales de Álgebra Lineal. El Capítulo 1 analiza vectores, matrices y sistemas de ecuaciones lineales, mientras el Capítulo 2 constituye una introducción a los determinantes.

Sin embargo, aún en estos primeros capítulos hay secciones opcionales que tienen el propósito de plantear algo más avanzado a los estudiantes. Por ejemplo, la Sección 1.11 contiene inversas unilaterales de matrices no cuadradas; en la Sección 2.3 se proporciona una demostración completa de que det AB =

det A det B. La prueba de este resultado primordial utilizando matrices elementales, generalmente no se incluye en textos introductorios.

Una característica importante del texto es la aparición frecuente del Teorema Resumen, que establece puntos de unión entre temas dispares en el estudio de matrices y transformaciones lineales. El Teorema aparece por primera vez en la Sección 1.2. Versiones más completas se tienen en las Secciones 1.8, 1.10, 2.5, 4.4, 4.7, 5.4 y 6.1.

El Capítulo 3 analiza vectores en el plano y en el espacio. Muchos de los temas de este capítulo se cubren en una secuencia de cálculo. Puesto que gran parte del Álgebra Lineal trata de espacios vectoriales abstractos, los estudiantes necesitan una buena provisión de ejemplos concretos, más fácilmente proporcionados por el estudio de vectores en el plano y en el espacio. Los Capítulos 4 y 5 que contienen temas más difíciles y abstractos se ilustran con ejemplos tomados del muy concreto Capítulo 3.

El Capítulo 4 contiene una introducción a los espacios vectoriales generales, y es necesariamente más abstracto que los capítulos que lo anteceden. Traté, sin embargo, de presentar el material como una extensión natural de las propiedades de los vectores en el plano —que es realmente como evoluciona el tema. Al final del Capítulo 4 se agrega una sección optativa en la cual demuestro que todo espacio vectorial tiene una base. En este proceso se analizan los conjuntos parcialmente ordenados y el lema de Zorn. Este material es necesariamente más abstracto que cualquier otro del libro y puede ser omitido. Sin embargo, creo firmemente que así como el Álgebra Lineal se considera como el primer curso de matemáticas en el cual las demostraciones son tan importantes como los cálculos, la demostración de este resultado fundamental debe ponerse al alcance de los estudiantes más motivados.

En el Capítulo 5 continúa la discusión comenzada en el Capítulo 4 con una introducción a las transformaciones lineales de un espacio vectorial a otro. El capítulo comienza con dos ejemplos que muestran cómo tales transformaciones pueden surgir de manera natural.

El Capítulo 6 describe la teoría de los valores y los vectores característicos (o propios). Su introducción se localiza en la Sección 6.1 y una aplicación biológica detallada se da en la Sección 6.2. En las Secciones 6.3, 6.4 y 6.5 interviene, en la diagonalización de una matriz, mientras que la Sección 6.6 ilustra, en unos cuantos casos, cómo puede ser reducida una matriz a una forma canónica de Jordan. En la Sección 6.8 presento dos de mis resultados favoritos de la teoría de matrices: el teorema de Cayley-Hamilton y el teorema del Círculo de Gershgorin. El último, escasamente discutido en un texto de Álgebra Lineal elemental, proporciona una manera más fácil de estimar los valores característicos de cualquier matriz.

En este Capítulo 6, tuve que tomar una difícil decisión: analizar los valores y vectores característicos completos o no hacerlo. Decidí incluirlos porque me pareció lo más indicado. Algunas de las matrices más "atractivas" tienen valores característicos complejos. Definir un valor característico como un número real solamente, puede hacer que las cosas parezcan más simples, pero esto ciertamente es erróneo. Más aún, en muchas aplicaciones concernientes a los valores característicos (incluyendo parte del material de la Sección 6.7) los modelos más interesantes dan por resultado fenómenos periódicos, y éstos incluyen valores

característicos complejos. Los números complejos no se evitan en este libro. Para los alumnos que no los han estudiado antes, las pocas propiedades que necesitan se describen ampliamente en el Apéndice 2.

En estos tiempos de calculadoras y computadoras al alcance de todos, es importante describir en un texto de álgebra lineal cómo se hacen los cálculos en la "vida real". En los primeros capítulos muchos cálculos se hicieron con calculadora. El Capítulo 7 contiene una introducción a varias técnicas numéricas utilizadas para resolver sistemas de ecuaciones y determinar valores y vectores característicos. La Sección 7.1 analiza algunos de los problemas que pueden surgir cuando se resuelven problemas por computadora. Las técnicas expuestas en las Secciones 7.1, 7.2 y 7.3 pueden estudiarse en cualquier momento después del Capítulo 1. El material de la Sección 7.4 (referente al cálculo de valores y vectores característicos) utiliza material presentado en la Sección 6.1.

Este libro contiene dos apéndices —uno sobre inducción matemática y otro acerca de los números complejos. Algunas de las demostraciones dadas en él utilizan inducción matemática, y para los estudiantes que no han manejado este importante método, el Apéndice 1 les proporciona una breve introducción.

Ahora una palabra sobre la dependencia recíproca de los capítulos. El libro fue escrito en forma secuencial, cada capítulo depende del precedente, con dos excepciones. El Capítulo 6 puede ser cubierto sin referirse a mucho del material del 5. El Capítulo 7 (excepto la Sección 7.4) puede ser tratado después del 1. Por otra parte, las secciones marcadas "opcional" se pueden omitir sin pérdida de continuidad.

En mi opinión, la parte más importante de cualquier texto de matemáticas elementales es el uso de ejemplos o problemas. Considero haber aprendido mucho de álgebra lineal resolviendo problemas, y creo que esto se verifica también para la mayoría de los estudiantes. Así, en muchas secciones he incluido tantos ejemplos como sea razonable, seguidos por un buen número de problemas de ejercicios y otros de naturaleza más teórica. Son complementados por ejercicios de repaso al final de cada capítulo. Los problemas más difíciles están marcados con una estrella (★) y unos pocos excepcionalmente difíciles con dos (★★).

Las respuestas de los problemas de número impar, incluyendo demostraciones, aparecen al final del libro. Las respuestas a los problemas de número par pueden obtenerse en un suplemento especial dirigiéndose al editor.

La numeración en el libro es la usual. Dentro de cada sección, los ejemplos, problemas, teoremas y ecuaciones están numerados consecutivamente a partir del 1. La referencia a un ejemplo, problema, teorema o ecuación fuera de la sección en la cual aparece, se formula por capítulo, sección y número. Así, el Ejemplo 4 de la Sección 2.5 se menciona simplemente, Ejemplo 4 en tal sección, pero fuera de la misma se refiere como Ejemplo 2.5.4.

Cambios notables en la tercera edición

Durante los siete años en que estuvieron publicadas las dos ediciones anteriores de *Álgebra Lineal* muchos lectores enviaron sus críticas y sugerencias. Más aún, algunas personas leyeron partes del manuscrito de la tercera edición y sugirieron

algunas formas de hacer el texto más accesible a los estudiantes. Por ello, he hecho un gran número de pequeños cambios que facilitarán la lectura de este libro.

Una de las críticas a la segunda edición fue la de que había una brecha muy grande entre el material analítico operacional de la primera parte del libro, y el material mucho más teórico que le seguía. Para no interrumpir la continuidad en el nivel del libro, he efectuado varios cambios notables incluyendo los siguientes:

- Los sistemas de ecuaciones y matrices aparecen ahora en un sólo capítulo. Algún material redundante fue eliminado.
- Las matrices elementales se presentan en la nueva Sección 1.10.
- En la Sección 1.11 se ha añadido una descripción de inversas unilaterales para matrices no cuadradas. Las nociones de mapeo o representación uno a uno y sobre se esbozan aquí.
- Una demostración completa de $\det AB = \det A \det B$ aparece ahora en la Sección 2.3.
- El contenido del Capítulo 3 (o sea el Capítulo 4 de la segunda edición) proporciona un resumen de geometría vectorial en $\mathbb{R}^2$ y $\mathbb{R}^3$. Este capítulo ha sido "flexibilizado" de modo que puede servir en dos formas: como de repaso para estudiantes que ya conocen el material por un curso de cálculo de varias variables, y como una breve introducción para aquellos estudiantes que no han tomado tal curso.
- En uno u otro casos el tema puede ser estudiado rápidamente para dar más tiempo al profesor en la exposición de los temas de espacio vectorial que siguen.
- La demostración de que todo espacio vectorial tiene una base se halla en la Sección 4.12.

Reconocimientos

Estoy muy agradecido con muchas personas cuya ayuda recibí durante el tiempo en que escribí este libro. Recibí una gran colaboración en la elaboración de problemas y conjuntos solución de parte de Robert Hollister, estudiante de posgrado de la Universidad de Montana, y de Robert Osterheld, también estudiante de posgrado en la Universidad de Purdue. Asimismo, agradezco a Academic Press, Inc., su autorización para utilizar material de mi libro *Calcular*, tercera edición, Capítulo 3. Tengo una deuda especial con Jim Harrison, supervisor editorial de matemáticas en Wadsworth, cuya dirección y conocimientos me ayudaron en las etapas más difíciles del proceso de escritura.

También quiero expresar mi gratitud a los siguientes revisores quienes hicieron aportaciones muy valiosas para la confiabilidad y didáctica de este texto.

Revisores de la tercera edición

John Chapman
Southwestern University

Thomas D. Forrest
Middle Tennessee State University

Henry Helson
University of California, Berkeley

William Jacob
Oregon State University

R. Bruce Lind
University of Puget Sound

Bryan Smith
University of Puget Sound

R. Van Enkevort
University of Puget Sound

Revisores de la segunda edición

P. A. Binding
University of Calgary

Joseph M. Cavanaugh
East Stroudsburg State College

Louis Florence
University of Toronto

Eugene Johnson
University of Iowa

Graham A. Chambers
University of Alberta

Ronald Eldringhoff
St. Louis Community College

John Sawka
University of Santa Clara

Keith Stumpff
Central Missouri State University

Revisores de la primera edición

Donald F. Bailey
Cornell College

Richard W. Ball
Auburn University

Sandra A. Bollinger
Longwood College

James Bradley
Roberts Wesleyan College

Joel V. Brawley
Clemson University

Paul Bugl
University of Hartford

Barbara Ann Greim
University of North Carolina en Wilmington

Charles H. Haggard
Transylvania University

David Kinsey
Indiana State University

Thomas Lupton
University of North Carolina en Wilmington

J. J. Malone
Worcester Polytechnic Institute

Gary Chartrand
Western Michigan University

Richard J. Easton
Indiana State University

Gary L. Eerkes
Gonzaga University

Garret J. Etgen
University of Houston

William R. Fuller
Purdue University

John D. Fulton
Clemson University

Lillian Gough
University of Wisconsin

Douglas D. Smith
Central Michigan University

T. W. Tucker
Dartmouth College

Marcellus E. Waddill
Wake Forrest University

James Wall
Auburn University

Edwin L. Marsden
Norwich University

Martin E. Nass
Iowa Central Community College

John Petro
Western Michigan University

William B. Rundberg
College of San Mateo

William Rundell
Texas A & M University

David Schedler
Virginia Commonwealth University

Jack C. Slay
Mississippi State University

Cary Webb
Chicago State University

James J. Woeppel
Indiana University Southeast

Dennis G. Zill
Loyola Marymount University

Contenido



Al estudiante

Grupo Editorial Iberoamérica *en su esfuerzo permanente de producir cada vez mejores textos, pone en tus manos esta nueva obra, en la que se ha puesto la más alta calidad en los aspectos teórico y didáctico, así como en diseño y presentación, con el objetivo de proporcionarte la mejor herramienta, no sólo para facilitarte el aprendizaje sino también para hacértelo más estimulante.*

Éste, como cualquiera de nuestros libros, ha sido cuidadosamente seleccionado para que encuentres en él un pilar de tu preparación, y un complemento ideal a la enseñanza del maestro. Lo didáctico de la presentación de sus temas hace que lo consideres el mejor auxiliar, y el que llevas a todas partes.

Lo anterior es parte de nuestro propósito de ser partícipes en una mejor preparación de profesionales, contribuyendo así a la urgente necesidad de un mayor desarrollo de nuestros países hispanohablantes.

Sabemos que esta obra será fundamental en tu biblioteca, y tal vez la más inmediata y permanente fuente de consulta.

Como uno de nuestros intereses principales es hacer mejores libros en equipo con profesores y estudiantes, agradeceremos tus comentarios y sugerencias o cualquier observación que contribuya al enriquecimiento de nuestras publicaciones.

Grupo Editorial Iberoamérica
. . . presente en tu formación profesional

CAPÍTULO 1

Sistemas de ecuaciones lineales y matrices

1.1 Introducción

Éste es un libro de álgebra lineal. Si buscamos la palabra "lineal" en un diccionario, encontraremos algo como lo siguiente: lineal, adj. Relativo a las líneas o de aspecto de línea.* En Matemáticas, la palabra "lineal" significa algo más que eso. Sin embargo, gran parte de la teoría del álgebra lineal elemental es de hecho una generalización de las propiedades de las líneas rectas. Como repaso, damos aquí algunos de los hechos fundamentales acerca de las citadas rectas:

i. La **pendiente** m de una recta que pasa por los puntos (x_1, y_1) y (x_2, y_2) está dada por (si $x_2 \neq x_1$).

$$m = \frac{y_2 - y_1}{x_2 - x_1} = \frac{\Delta y}{\Delta x}$$

ii. Si $x_2 - x_1 = 0$ y $y_2 \neq y_1$ entonces la recta es vertical y se dice que la pendiente no está definida.†

iii. Cualquier recta (excepto una con pendiente indefinida), se puede describir expresando su ecuación en la forma simplificada $y = mx + b$, donde m es la pendiente de la recta y b es la ordenada al origen (el valor de y en el punto donde la recta cruza al eje y).

iv. Dos rectas son paralelas si y sólo si tienen la misma pendiente.

v. Si la ecuación de una recta es $ax + by = c$ $(b \neq 0)$ entonces, como se ve fácilmente, $m = -a/b$.

vi. Si m_1 es la pendiente de la recta L_1, m_2 la de L_2, $m_1 \neq 0$ y L_1, L_2 son perpendiculares, entonces $m_2 = -1m_1$.

---

* (N. del E.) Tomado del Diccionario Larousse Universal Ilustrado, 1968.

† A veces se dice que una recta vertical "tiene pendiente infinita" o que "no tiene pendiente".

vii. Las rectas paralelas al eje x tienen pendiente cero.
viii. Las rectas paralelas al eje y tienen pendiente indefinida.

En la siguiente sección ilustraremos las relaciones entre la solución de sistemas de ecuaciones y el modo de encontrar los puntos de intersección de pares de líneas rectas.

1.2 Dos ecuaciones lineales en dos incógnitas

Consideremos el siguiente sistema de dos ecuaciones lineales en dos incógnitas x_1 y x_2:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1 \\ a_{21}x_1 + a_{22}x_2 &= b_2 \end{aligned} \tag{1}$$

donde a_{11}, a_{12}, a_{21}, a_{22}, b_1 y b_2 son números dados. Cada una de estas ecuaciones es la ecuación de una línea recta (en el plano $x_1 x_2$ en vez del plano xy). La pendiente de la primera recta es $-a_{11}/a_{12}$ y la pendiente de la segunda es $-a_{21}/a_{22}$ (si $a_{12} \neq 0$ y $a_{22} \neq 0$). Una *solución* del sistema (1) es un par de números, denotados (x_1, x_2), que satisface (1). Las preguntas que surgen naturalmente son: ¿cuándo (1) tiene soluciones? y, si las tiene, ¿cuántas son? Responderemos a estas preguntas después de ver algunos ejemplos. En estos ejemplos usaremos dos propiedades importantes del álgebra elemental:

Propiedad A. Si $a = b$ y $c = d$ entonces $a + c = b + d$.

Propiedad B. Si $a = b$ y c es cualquier número real, entonces $ca = cb$.

La Propiedad A dice que si sumamos dos ecuaciones obtenemos una tercera ecuación válida. La B dice que si multiplicamos ambos lados de una ecuación por una constante obtenemos una segunda ecuación válida.

Ejemplo 1 Considere el sistema

$$\begin{aligned} x_1 - x_2 &= 7 \\ x_1 + x_2 &= 5 \end{aligned} \tag{2}$$

Si sumamos las dos ecuaciones tenemos, por la propiedad A, la siguiente ecuación: $2x_1 = 12$ (o $x_1 = 6$). Entonces, de la segunda ecuación tenemos $x_2 = 5 - x_1 = 5 - 6 = -1$. Así, el par $(6, -1)$ satisface el sistema (2) y por la forma en que encontramos la solución vemos que es el único par de números que lo hace. Esto es, el sistema (2) tiene *solución única*.

Ejemplo 2 Considere el sistema

$$\begin{aligned} x_1 - x_2 &= 7 \\ 2x_1 - 2x_2 &= 14 \end{aligned} \tag{3}$$

Es claro que estas dos ecuaciones son equivalentes. Para ver esto multiplicamos la primera por 2. (Esto es válido por la Propiedad B.) Entonces, $x_1 - x_2 = 7$ ó $x_2 = x_1 - 7$. Así, el par $(x_1, x_1 - 7)$ es una solución del sistema (3) para cualquier número real x_1. Esto es, el sistema (3) tiene un *número infinito de soluciones*. Por ejemplo, los siguientes pares son soluciones: (7, 0), (0, −7), (8, 1), (1, −6), (3, −4) y (−2, −9).

Ejemplo 3 Considere el sistema

$$\begin{aligned} x_1 - x_2 &= 7 \\ 2x_1 - 2x_2 &= 13 \end{aligned} \tag{4}$$

Al multiplicar la primera ecuación por 2 (lo cual está permitido por la Propiedad B) nos da $2x_1 - 2x_2 = 14$. Esto contradice a la segunda ecuación. Así, el sistema (4) *no tiene solución*.

Es fácil de explicar geométricamente lo que sucede en los ejemplos anteriores. Primero, recordemos que las ecuaciones del sistema (1) son ecuaciones de rectas. Una solución de (1) es un punto (x_1, x_2) que está en ambas rectas. Si las dos rectas no son paralelas, se intersecan en un solo punto; si son paralelas nunca se intersecan (no tienen puntos en común), o son la misma recta (tienen un número infinito de puntos en común). En el Ejemplo 1 las rectas tienen pendientes 1 y −1, respectivamente. Por lo tanto, no son paralelas. Sólo poseen el punto (6, −1) en común. En el Ejemplo 2 las rectas son paralelas (pendiente igual a 1) y coincidentes. En el Ejemplo 3 las rectas son paralelas y distintas. Estas relaciones se ilustran en la Figura 1.1.

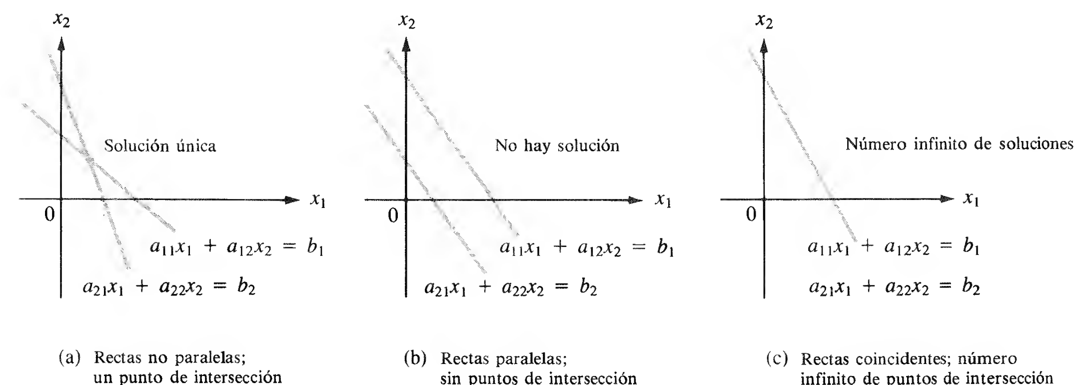


Figura 1.1

Ahora resolvamos el sistema (1) formalmente. Tenemos

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1 \\ a_{21}x_1 + a_{22}x_2 &= b_2 \end{aligned}$$

Si multiplicamos la primera ecuación por a_{22} y la segunda por a_{12} tenemos

$$\begin{aligned} a_{11}a_{22}x_1 + a_{12}a_{22}x_2 &= a_{22}b_1 \\ a_{12}a_{21}x_1 + a_{12}a_{22}x_2 &= a_{12}b_2 \end{aligned} \tag{5}$$

Antes de continuar notemos que el sistema (1) y el sistema (5) son *equivalentes*. Esto significa que cualquier solución del sistema (1) es una solución del sistema (5) y viceversa. Ello se sigue inmediatamente de la Propiedad B. Después restamos la segunda ecuación de la primera y obtenemos

$$(a_{11}a_{22} - a_{12}a_{21})x_1 = a_{22}b_1 - a_{12}b_2 \tag{6}$$

Observemos que si $a_{11}a_{22} - a_{12}a_{21} \neq 0$, entonces podemos dividir entre esta cantidad y tener

$$x_1 = \frac{a_{22}b_1 - a_{12}b_2}{a_{11}a_{22} - a_{12}a_{21}}$$

Ahora es posible insertar este valor de x_1 en el sistema (1), despejar x_2 y habremos encontrado la solución única del sistema. Definimos el *determinante* del sistema (1) como

Determinante del sistema $(1) = a_{11}a_{22} - a_{12}a_{21}$ (7)

y habremos mostrado lo siguiente:

Si el determinante del sistema (1) $\neq 0$, entonces el sistema tiene una única solución. (8)

¿Cómo se relaciona esta afirmación con lo que hemos discutido anteriormente? En el sistema (1) vemos que la pendiente de la primera recta es $-a_{11}/a_{12}$ y la pendiente de la segunda es $-a_{21}a_{22}$.* En los Problemas 31, 32 y 33 se pide mostrar que el determinante del sistema (1) es cero si y sólo si las rectas son paralelas (tienen la misma pendiente). Así, si el determinante *no* es cero, las rectas no son paralelas y el sistema tiene una solución única.

Ahora expresaremos en un teorema los hechos discutidos anteriormente. Es un teorema que será generalizado en secciones posteriores de este capítulo y en capítulos siguientes. Este teorema, llamado "Teorema Resumen", nos dará una indicación de nuestro progreso en los capítulos siguientes. Cuando todas sus partes hayan sido probadas, veremos una relación notable entre varios conceptos importantes del álgebra lineal.

* Estamos suponiendo que ni a_{12} ni a_{22} son cero. Si a_{12} o a_{22} fueran cero entonces el sistema (1) sería fácilmente resuelto. Si $a_{12} = 0$, por ejemplo, entonces $x_1 = b_1/a_{11}$. Por tanto no nos preocuparemos de esta posibilidad.

Teorema 1 ***Teorema Resumen, Versión 1.*** El sistema

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1 \\ a_{21}x_1 + a_{22}x_2 &= b_2 \end{aligned}$$

de dos ecuaciones en las incógnitas x_1 y x_2 no tiene solución, tiene una única solución o tiene un número infinito de soluciones. Tendrá:

i. Una única solución si y sólo si su determinante es distinto de cero.
ii. Ninguna solución o un número infinito de soluciones si y sólo si su determinante es cero.

En la Sección 1.6 discutiremos sistemas de m ecuaciones en n incógnitas y veremos que siempre existe una solución, ninguna o un número infinito de soluciones. En el Capítulo 2 definiremos y calcularemos determinantes para sistemas de n ecuaciones en n incógnitas y encontraremos que nuestro Teorema Resumen (Versión 1) es cierto en este contexto general.

Problemas 1.2

En los Problemas del 1 *al* 12, *encuentre todas las soluciones (si existen) a los sistemas dados. En cada caso calcule el determinante.*

1. $x_1 - 3x_2 = 4$; $-4x_1 + 2x_2 = 6$
2. $2x_1 - x_2 = -3$; $5x_1 + 7x_2 = 4$
3. $2x_1 - 8x_2 = 5$; $-3x_1 + 12x_2 = 8$
4. $2x_1 - 8x_2 = 6$; $-3x_1 + 12x_2 = -9$
5. $6x_1 + x_2 = 3$; $-4x_1 - x_2 = 8$
6. $3x_1 + x_2 = 0$; $2x_1 - 3x_2 = 0$
7. $4x_1 - 6x_2 = 0$; $-2x_1 + 3x_2 = 0$
8. $5x_1 + 2x_2 = 3$; $2x_1 + 5x_2 = 3$
9. $2x_1 + 3x_2 = 4$; $3x_1 + 4x_2 = 5$
10. $ax_1 + bx_2 = c$; $ax_1 - bx_2 = c$
11. $ax_1 + bx_2 = c$; $bx_1 + ax_2 = c$
12. $ax_1 - bx_2 = c$; $bx_1 + ax_2 = d$

13. Encuentre condiciones para a y b de forma que el sistema del Problema 10 tenga una única solución.

14. Encuentre condiciones para a, b y c de forma que el sistema del Problema 11 tenga un número infinito de soluciones.

15. Encuentre condiciones para a, b, c y d de forma que el sistema del Problema 12 no tenga soluciones.

En los Problemas del 16 *al* 21, *encuentre el punto de intersección (si existe alguno) de las dos rectas.*

16. $x - y = 7$; $2x + 3y = 1$
17. $y - 2x = 4$; $4x - 2y = 6$
18. $4x - 6y = 7$; $6x - 9y = 12$
19. $4x - 6y = 10$; $6x - 9y = 15$
20. $3x + y = 4$; $y - 5x = 2$
21. $3x + 4y = 5$; $6x - 7y = 8$

Sea L una recta y sea $L_\perp$ la recta perpendicular a L que pasa por un punto dado P. La distancia *de L a P se define como la distancia* entre P y el punto de*

* Recuerde que si (x_1, y_1) y (x_2, y_2) son dos puntos en el plano xy, entonces la distancia d entre estos dos puntos está dada por $d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$.

intersección de L y $L_\perp$. En los Problemas del 22 al 27 encuentre la distancia entre la recta y el punto dados.

22. $x-y=6$; $(0, 0)$
23. $2x+3y=-1$; $(0, 0)$
24. $3x+y=7$; $(1, 2)$
25. $5x-6y=3$; $(2, \frac{16}{5})$
26. $2y-5x=-2$; $(5, -3)$
27. $6y+3x=3$; $(8, -1)$
28. Encuentre la distancia entre la recta $2x-y=6$ y el punto de intersección de las rectas $2x-3y=1$ y $3x+6y=12$.

★ 29. Demuestre que la distancia entre el punto (x_1, y_1) y la recta $ax+by=c$ está dada por

$$d=\frac{|ax_1+by_1-c|}{\sqrt{a^2+b^2}}$$

30. Un zoológico tiene aves (bípedos) y bestias (cuadrúpedos). Si el zoológico tiene 60 cabezas y 200 patas, ¿cuántas aves y cuántas bestias viven allí?
31. Suponga que el determinante del sistema (1) es cero. Muestre que las rectas dadas en (1) son paralelas.
32. Si existe una única solución al sistema (1) muestre que su determinante es distinto de cero.
33. Si el determinante del sistema (1) es distinto de cero, demuestre que el sistema tiene una única solución.
33. Si el determinante del sistema (1) es distinto de cro, demuestre que el sistema tiene una única solución.
34. Una fábrica de porcelana manufactura tazas y platos de cerámica. Por cada taza o plato, un trabajador mide una cantidad fija de material y la coloca en una máquina moldeadora, de la cual sale automáticamente vidriada y cocida. En promedio, un trabajador necesita 3 minutos para iniciar el proceso en el caso de una taza, y 2 minutos para un plato. El material para la taza cuesta \$25 y para el plato \$20. Si se asignan \$4,400 diarios para producción de tazas y platos, ¿cuántas unidades de cada uno de estos productos se pueden manufacturar en una jornada de 8 horas, si un trabajador labora cada minuto del día y exactamente se gastan los \$4,400 en los materiales?
35. Conteste la pregunta del Problema 34 si los materiales para una taza y un plato cuestan \$15 y \$10, respectivamente, y se gastan \$2,400 en una jornada de 8 horas.
36. Conteste la pregunta del Problema 35 si se gastan \$2,500 en una jornada de 8 horas.
37. Una heladería vende solo helado con soda y leches malteadas. En el primero se usan 1 onza de jarabe y 4 onzas de helado. En la segunda, se utilizan 1 onza de jarabe y 3 onzas de helado. Si el expendio usa 4 galones de helado y 5 cuartos de jarabe en un día, ¿cuántos helados con soda y malteadas vende diariamente? [Equivalencias: 1 cuarto = 32 onzas; 1 galón = 128 onzas.]

1.3 Vectores

El estudio de los vectores y matrices es parte medular del álgebra lineal. El estudio de vectores comenzó esencialmente con el trabajo del gran matemático irlandés William Rowan Hamilton (1805-1865).* Su deseo de encontrar un modo de representar ciertos objetos en el plano y en el espacio, lo llevó al descubrimiento de lo que llamó *cuaterniones*. Este concepto lo condujo al de lo

* Véase la reseña biográfica.

que actualmente se denomina *vectores*. Durante la vida de Hamilton, y en lo que restó del siglo diecinueve, existió un considerable debate sobre la utilidad de los cuaterniones o cuaternios y de los vectores. A fin de siglo, el gran físico británico Lord Kelvin, escribió acerca de los cuaternios: "aun cuando son notablemente ingeniosos, han demostrado ser una mala fortuna para todos aquellos que de alguna manera los han estudiado [así como] los vectores. . . nunca han sido de la más remota utilidad a criatura alguna".

Sin embargo Kelvin estaba equivocado. Actualmente, casi todas las áreas de la física clásica y moderna son representadas por medio del lenguaje de los vectores. También se usan cada vez más frecuentemente en las ciencias biológicas y sociales.*

En la Sección 1.2 describimos la solución a un sistema de dos ecuaciones en dos incógnitas como un par de números (x_1, x_2). En el Ejemplo 1.6.1 expresaremos la solución de un sistema de tres ecuaciones en tres incógnitas como la terna de números $(4, -2, 3)$. Ambos, (x_1, x_2) y $(4, -2, 3)$ son *vectores*.

Vector renglón de n componentes

Definimos un *vector renglón de n componentes* (o *n-dimensional*) como un conjunto *ordenado* de n números escrito como

$$(x_1, x_2, \ldots, x_n) \tag{1}$$

Vector columna de n componentes

Un *vector columna de n componentes* (o *n-dimensional*) es un conjunto *ordenado* de n números escrito como

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \tag{2}$$

En (1) o (2), x_1 se llama la *primera componente* del vector, x_2 es la *segunda componente*, y así sucesivamente. En general, x_k es la *k-ésima componente* del vector.

Por simplicidad, frecuentemente nos referimos a un vector renglón n-dimensional como un *vector renglón* o un *n-vector*. De igual manera, usamos el término *vector columna* (o n-vector) para denotar a un vector columna n-dimensional. Cualquier vector con todas sus componentes iguales a cero se llama *vector cero*.

Ejemplo 1 Los siguientes son ejemplos de vectores:

i. $(3, 6)$ es un vector renglón con dos componentes.

ii. $\begin{pmatrix} 2 \\ -1 \\ 5 \end{pmatrix}$ es un vector columna con tres componentes.

* Descripciones interesantes del desarrollo del análisis vectorial moderno pueden verse en el libro de M.J. Crowe, *A History of Vector Analysis* (Notre Dame: University of Notre Dame Press, 1967) o en el excelente libro de Morris Kline, *Mathematical Thought from Ancient to Modern Times* (Nueva York: Oxford University Press, 1972), Cap. 32.

iii. es un vector renglón con cuatro componentes.

iv. $\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$ es un vector columna y un vector cero.

Advertencia. En la definición de vector, la palabra "ordenado" es esencial. Dos vectores con las mismas componentes escritas en diferente orden *no* son los mismos. Así, por ejemplo, los vectores renglón (1,2) y (2,1) no son iguales.

De aquí en adelante denotaremos los vectores con letras minúsculas en tipo negro como **u**, **v**, **a**, **b**, **c**, etc. El vector cero se denota por **0**.

Los vectores surgen de diferentes maneras. Suponga que el comprador de una planta manufacturera debe ordenar cantidades diferentes de acero, aluminio, aceite y papel. Puede anotar las cantidades ordenadas con un simple vector. El vector $\begin{pmatrix} 10 \\ 30 \\ 15 \\ 60 \end{pmatrix}$ indica que se han ordenado 10 unidades de acero, 30 de aluminio, y así sucesivamente.

Observación. Aquí vemos por qué es importante el orden en el que son escritas las componentes de un vector. Es claro que los vectores $\begin{pmatrix} 30 \\ 15 \\ 60 \\ 10 \end{pmatrix}$ y $\begin{pmatrix} 10 \\ 30 \\ 15 \\ 60 \end{pmatrix}$ significan cosas muy diferentes para el comprador.

Ahora describiremos algunas propiedades de los vectores. Como sería repetitivo hacerlo primero para vectores renglón y después para vectores columna, daremos todas las definiciones en términos de vectores columna. Las definiciones para los vectores renglón son similares.

Las componentes de todos los vectores en este libro son números reales o complejos.* Usamos el símbolo $\mathbb{R}^n$ para denotar el conjunto de todos los n-vectores $\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$, donde cada a_i es un número real. De igual manera, usamos

* Un número complejo es un número de la forma $a + ib$, donde a y b son números reales e $i = \sqrt{-1}$. Una descripción de los números complejos se da en el Apéndice 2. No encontraremos vectores complejos otra vez hasta el Capítulo 4; serán de especial utilidad en el Capítulo 6. Por tanto, mientras no se diga otra cosa suponemos que todos los vectores tienen componentes reales.

Sir William Rowan Hamilton, 1805-1865



Sir William Rowan Hamilton (*The Granger Collection*)

Nacido en Dublín en 1805, en donde pasó la mayor parte de su vida, William Rowan Hamilton fue, sin la menor duda, el más grande matemático irlandés de su época. Su padre (un abogado) y su madre murieron cuando era aún pequeño. Su tío, un lingüista, se encargó de la educación del muchacho. Cuando cumplió cinco años, Hamilton podía leer inglés, hebreo, latín y griego. Cuando tenía trece años, no sólo dominaba todas las lenguas del continente europeo, sino también el sánscrito, chino, persa, árabe, malayo, indostánico, bengalí y algunos otros más. A Hamilton le gustaba escribir poesía, tanto de niño como de adulto, y entre sus amigos se encontraban algunos grandes poetas ingleses como Samuel Taylor Coleridge y William Wordsworth. La poesía de Hamilton se consideraba tan mala que resulta una fortuna que haya desarrollado otros intereses —especialmente en matemáticas.

Aunque Hamilton gustó de las matemáticas desde muy joven, su interés por las mismas se acrecentó cuando, a los quince años, por casualidad, conoció al velocísimo calculista norteamericano Zerah Colburn. Poco después Hamilton comenzó a leer importantes libros de matemáticas de la época. En 1823, a los dieciocho años de edad, descubrió un error en la *Mecánique céleste* de Simón Laplace y escribió un impresionante artículo sobre el tema. Un año después, ingresó al Trinity College en Dublín.

La carrera universitaria de Hamilton fue asombrosa. A los 21 años, cuando todavía no se graduaba, había impresionado tanto al cuerpo docente, que fue nombrado Astrónomo Real de Irlanda y catedrático de Astronomía de la Universidad. Al poco tiempo de este suceso, escribió lo que ahora se considera un trabajo clásico de Óptica. Utilizando sólo teoría matemática predijo refracción cónica en cierto tipo de cristales. Posteriormente, su teoría fue confirmada por los físicos. Con base principalmente en su trabajo, se concedió a Hamilton un título nobiliario en 1835.

Su primer gran artículo puramente matemático apareció en 1833. En este trabajo, describió un modo algebraico para manipular parejas de números reales. Este estudio da reglas que se usan ahora para sumar, restar, multiplicar y dividir los números complejos. Al principio, sin embargo, Hamilton no pudo hallar un método para la multiplicación de tríadas o n-adas de números cuando $n > 2$. Diez años dedicó a pensar en este problema, y se dice que lo resolvió por inspiración mientras caminaba un día sobre el puente de Brougham en Dublín en 1843. La clave fue la de descartar la bien conocida propiedad conmutativa de la multiplicación. Los nuevos conceptos se llamaron *cuaterniones* o *cuaternios*, y fueron los precursores de lo que ahora llamamos *vectores*.

Durante el resto de su vida, Hamilton dedicó la mayor parte de su tiempo desarrollando el álgebra de los cuaternios. Creía que habrían de tener un significado revolucionario en la física matemática. Su monumental obra, *Treatise on quaternions* fue publicada en 1853. Después de esto, trabajó en una obra mayor, *Elements of quaternions*. Aun cuando Hamilton murió en 1865 antes de terminar sus "Elementos", el trabajo fue publicado por su hijo en 1866.

Los estudiantes de matemáticas y de física conocen a Hamilton en una gran variedad de contextos. En física matemática, por ejemplo, se estudia la función hamiltoniana, que frecuentemente representa la energía total en un sistema, así como las ecuaciones diferenciales de Hamilton-Jacobi. En la teoría de matrices, el teorema de Cayley-Hamilton, establece que toda matriz satisface su propia ecuación característica.

A pesar del gran trabajo que Hamilton estaba desarrollando, sus últimos años fueron tormentosos. Su esposa estaba semiinválida, y él se convirtió en alcohólico. Por lo tanto, es confortante destacar que durante esos últimos años, la naciente Academia Nacional de Ciencias de Estados Unidos, eligió a Sir William Rowan Hamilton como su primer miembro extranjero.

el símbolo $\mathbb{C}^n$ para denotar al conjunto de todos los n-vectores $\begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$, donde cada c_i es un número complejo. En el Capítulo 3 discutiremos los conjuntos $\mathbb{R}^2$ (vectores en el plano) y $\mathbb{R}^3$ (vectores en el espacio). En el Capítulo 4 estudiaremos conjuntos arbitrarios de vectores.

Definición 1 ***Igualdad de vectores.*** Dos vectores columna (o renglón) **a** y **b** son *iguales* si y sólo si* tienen el mismo número de componentes y sus componentes correspondientes son iguales. En símbolos, los vectores $\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$ y $\mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$ son iguales si y sólo si $a_1 = b_1, a_2 = b_2, \ldots, a_n = b_n$.

Definición 2 ***Suma de vectores.*** Sean $\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$ y $\mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$ n-vectores. Entonces la suma de **a** y **b** se define como

$$\mathbf{a}+\mathbf{b} = \begin{pmatrix} a_1+b_1 \\ a_2+b_2 \\ \vdots \\ a_n+b_n \end{pmatrix} \tag{3}$$

Ejemplo 2 $\begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix} + \begin{pmatrix} -6 \\ 7 \\ 5 \end{pmatrix} = \begin{pmatrix} -5 \\ 9 \\ 9 \end{pmatrix}$

Ejemplo 3 $\begin{pmatrix} 2 \\ -1 \end{pmatrix} + \begin{pmatrix} -2 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

* El término "si y sólo si" se aplica a dos proposiciones: "La proposición A si y sólo si la proposición B" significa que las proposiciones A y B son equivalentes. Esto es, si la proposición A es cierta entonces la proposición B es cierta y si la proposición B es cierta entonces la proposición A es cierta. Dicho de otra forma, significa que no se puede tener una sin la otra.

Advertencia. Es esencial que **a** y **b** tengan el mismo número de componentes. Por ejemplo, la suma $\begin{pmatrix} 2 \\ 3 \end{pmatrix} + \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ no está definida pues los vectores con dos o tres componentes son diferentes clases de objetos y no se pueden sumar. Más aún, no es posible sumar un vector renglón y un vector columna. Por ejemplo, la suma $\begin{pmatrix} 1 \\ 2 \end{pmatrix} + (3, 5)$ *no* está definida.

Cuando trabajamos con vectores, nos referimos a los números como *escalares* (que pueden ser reales o complejos, dependiendo de que los vectores en cuestión sean reales o complejos).*

Definición 3 ***Multiplicación de vectores por un escalar.*** Sea $\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$ un vector y α un escalar. Entonces el producto $\alpha\mathbf{a}$ está dado por

$$\alpha\mathbf{a} = \begin{pmatrix} \alpha a_1 \\ \alpha a_2 \\ \vdots \\ \alpha a_n \end{pmatrix} \tag{4}$$

Esto es, al multiplicar un vector por un escalar simplemente multiplicamos cada componente del vector por el escalar.

Ejemplo 4 $3\begin{pmatrix} 2 \\ -1 \\ 4 \end{pmatrix} = \begin{pmatrix} 6 \\ -3 \\ 12 \end{pmatrix}$

* *Nota histórica:* El término "escalar" se originó con Hamilton. Su definición de cuaternión incluía lo que él llamó una "parte real" y una "parte imaginaria". En su artículo "On Quaternions, or on a New System of Imaginaries in Algebra," en *Philosophical Magazine*, tercera serie, 25(1844): 26-27, escribió: "La parte algebraicamente *real* puede recibir... todos los valores contenidos en una *escala* de progresión de números del infinito negativo al infinito positivo; nosotros la llamaremos, por tanto, la *parte escalar* o, simplemente, el *escalar* del cuaternión. . . " En el mismo artículo Hamilton definió la parte imaginaria de su cuaternión como la parte *vectorial*. Aunque no fue la primera vez que se usó la palabra "vector", fue la primera vez que se usó en el contexto de las definiciones de esta sección. Se puede decir que el artículo del cual se tomó la nota precedente marca el principio del análisis vectorial moderno.

Nota. Juntando la Definición 1 y la Definición 2 podemos definir la diferencia de dos vectores como

$$\mathbf{a}-\mathbf{b}=\mathbf{a}+(-1)\mathbf{b} \qquad (5)$$

Esto significa que si $\mathbf{a}=\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$ y $\mathbf{b}=\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$, entonces $\mathbf{a}-\mathbf{b}=\begin{pmatrix} a_1-b_1 \\ a_2-b_2 \\ \vdots \\ a_n-b_n \end{pmatrix}$

Ejemplo 5 Sean $\mathbf{a}=\begin{pmatrix} 4 \\ 6 \\ 1 \\ 3 \end{pmatrix}$ y $\mathbf{b}=\begin{pmatrix} -2 \\ 4 \\ -3 \\ 0 \end{pmatrix}$. Calcule $2\mathbf{a}-3\mathbf{b}$.

Solución

$$2\mathbf{a}-3\mathbf{b}=2\begin{pmatrix} 4 \\ 6 \\ 1 \\ 3 \end{pmatrix}+(-3)\begin{pmatrix} -2 \\ 4 \\ -3 \\ 0 \end{pmatrix}=\begin{pmatrix} 8 \\ 12 \\ 2 \\ 6 \end{pmatrix}+\begin{pmatrix} 6 \\ -12 \\ 9 \\ 0 \end{pmatrix}=\begin{pmatrix} 14 \\ 0 \\ 11 \\ 6 \end{pmatrix}$$

Una vez visto cómo sumar vectores y cómo multiplicarlos por escalares, podemos demostrar algunas propiedades relativas a estas operaciones. Varias de estas propiedades se dan en el Teorema 1. Probaremos las partes (*ii*) y (*iii*) y dejaremos las partes restantes como ejercicios (vea los Problemas del 21 al 23).

Teorema 1 Sean **a**, **b** y **c** n-vectores y sean α y β escalares. Entonces:

i. $\mathbf{a}+\mathbf{0}=\mathbf{a}$
ii. $0\mathbf{a}=\mathbf{0}$ (Nótese que el cero de la izquierda es el número cero y el de la derecha es el vector nulo.)
iii. $\mathbf{a}+\mathbf{b}=\mathbf{b}+\mathbf{a}$ **(ley conmutativa)**
iv. $(\mathbf{a}+\mathbf{b})+\mathbf{c}=\mathbf{a}+(\mathbf{b}+\mathbf{c})$ **(ley asociativa)**
v. $\alpha(\mathbf{a}+\mathbf{b})=\alpha\mathbf{a}+\alpha\mathbf{b}$ **(ley distributiva para la multiplicación por un escalar)**
vi. $(\alpha+\beta)\mathbf{a}=\alpha\mathbf{a}+\beta\mathbf{a}$
vii. $(\alpha\beta)\mathbf{a}=\alpha(\beta\mathbf{a})$

Demostración de (ii) y (iii)

ii. Si $\mathbf{a}=\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$, entonces $0\mathbf{a}=0\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}=\begin{pmatrix} 0\cdot a_1 \\ 0\cdot a_2 \\ \vdots \\ 0\cdot a_n \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}=\mathbf{0}$.

iii. Sea $\mathbf{b}=\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$. Entonces $\mathbf{a}+\mathbf{b}=\begin{pmatrix} a_1+b_1 \\ a_2+b_2 \\ \vdots \\ a_n+b_n \end{pmatrix}=\begin{pmatrix} b_1+a_1 \\ b_2+a_2 \\ \vdots \\ b_n+a_n \end{pmatrix}=\mathbf{b}+\mathbf{a}$.

Aquí hemos usado el hecho de que para cualquier par de números x y y se tiene que $x+y=y+x$ y $0\cdot x=0$. ■

Ejemplo 6 Para ilustrar la ley asociativa vemos que

$$\left[\begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix}+\begin{pmatrix} -2 \\ 4 \\ -1 \end{pmatrix}\right]+\begin{pmatrix} 6 \\ -3 \\ 5 \end{pmatrix}=\begin{pmatrix} 1 \\ 5 \\ 1 \end{pmatrix}+\begin{pmatrix} 6 \\ -3 \\ 5 \end{pmatrix}=\begin{pmatrix} 7 \\ 2 \\ 6 \end{pmatrix}$$

mientras que

$$\begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix}+\left[\begin{pmatrix} -2 \\ 4 \\ -1 \end{pmatrix}+\begin{pmatrix} 6 \\ -3 \\ 5 \end{pmatrix}\right]=\begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix}+\begin{pmatrix} 4 \\ 1 \\ 4 \end{pmatrix}=\begin{pmatrix} 7 \\ 2 \\ 6 \end{pmatrix}$$

El Ejemplo 6 ilustra la importancia de la ley asociativa para la suma de vectores, puesto que si queremos sumar más de dos vectores podemos hacerlo solamente sumando dos de ellos a la vez. La ley asociativa nos dice que podemos hacer esto de dos diferentes maneras y llegar al mismo resultado. Si éste no fuera el caso, la suma de tres o más vectores sería más difícil de definir pues tendríamos que especificar cuándo queremos $(\mathbf{a}+\mathbf{b})+\mathbf{c}$ y cuándo $\mathbf{a}+(\mathbf{b}+\mathbf{c})$ para la suma $\mathbf{a}+\mathbf{b}+\mathbf{c}$.

Problemas 1.3

En los Problemas del 1 *al* 10 *efectúe la operación indicada con* $\mathbf{a}=\begin{pmatrix} -3 \\ 1 \\ 4 \end{pmatrix}$, $\mathbf{b}=\begin{pmatrix} 5 \\ -4 \\ 7 \end{pmatrix}$ *y* $\mathbf{c}=\begin{pmatrix} 2 \\ 0 \\ -2 \end{pmatrix}$

1. $\mathbf{a}+\mathbf{b}$
2. $3\mathbf{b}$
3. $-2\mathbf{c}$
4. $\mathbf{b}+3\mathbf{c}$
5. $2\mathbf{a}-5\mathbf{b}$
6. $-3\mathbf{b}+2\mathbf{c}$
7. $0\mathbf{c}$
8. $\mathbf{a}+\mathbf{b}+\mathbf{c}$
9. $3\mathbf{a}-2\mathbf{b}+4\mathbf{c}$
10. $3\mathbf{b}-7\mathbf{c}+2\mathbf{a}$

En los Problemas del 11 *al* 20 *efectúe la operación indicada con* $\mathbf{a}=(3,-1,4,2)$, $\mathbf{b}=(6,0,-1,4)$ *y* $\mathbf{c}=(-2,3,1,5)$. *Desde luego, primero es necesario extender las definiciones de esta sección a vectores renglón.*

11. $\mathbf{a}+\mathbf{c}$
12. $\mathbf{b}-\mathbf{a}$
13. $4\mathbf{c}$
14. $-2\mathbf{b}$
15. $2\mathbf{a}-\mathbf{c}$
16. $4\mathbf{b}-7\mathbf{a}$
17. $\mathbf{a}+\mathbf{b}+\mathbf{c}$
18. $\mathbf{c}-\mathbf{b}+2\mathbf{a}$
19. $3\mathbf{a}-2\mathbf{b}+4\mathbf{c}$
20. $\alpha\mathbf{a}+\beta\mathbf{b}+\gamma\mathbf{c}$

21. Sea $\mathbf{a}=\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix}$ y sea $\mathbf{0}$ el vector columna n-dimensional cero. Use las Definiciones 2 y 3 para mostrar que $\mathbf{a}+\mathbf{0}=\mathbf{a}$ y $0\mathbf{a}=\mathbf{0}$.
22. Sean $\mathbf{a}=\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix}$, $\mathbf{b}=\begin{pmatrix}b_1\\b_2\\\vdots\\b_n\end{pmatrix}$ y $\mathbf{c}=\begin{pmatrix}c_1\\c_2\\\vdots\\c_n\end{pmatrix}$. Calcule $(\mathbf{a}+\mathbf{b})+\mathbf{c}$ y $\mathbf{a}+(\mathbf{b}+\mathbf{c})$ y muestre que son iguales.
23. Sean $\mathbf{a}$ y $\mathbf{b}$ como en el Problema 22 y sean α y β escalares. Calcule $\alpha(\mathbf{a}+\mathbf{b})$ y $\alpha\mathbf{a}+\alpha\mathbf{b}$ y muestre que son iguales. De igual manera, calcule $(\alpha+\beta)\mathbf{a}$ y $\alpha\mathbf{a}+\beta\mathbf{a}$ y muestre que son iguales. Finalmente, muestre que $(\alpha\beta)\mathbf{a}=\alpha(\beta\mathbf{a})$.
24. Encuentre números α, β y γ tales que $(2,-1,4)+(\alpha,\beta,\gamma)=\mathbf{0}$.
25. En la fabricación de cierto producto se necesitan cuatro materias primas. El vector $\mathbf{d}=\begin{pmatrix}d_1\\d_2\\d_3\\d_4\end{pmatrix}$ representa una demanda dada para cada uno de los cuatro materiales para elaborar una unidad de su producto. Si $\mathbf{d}_1$ es el vector de demanda para la fábrica 1 y d_2 es el vector de demanda para la fábrica 2, ¿qué representan los vectores $\mathbf{d}_1+\mathbf{d}_2$ y $2\mathbf{d}_1$?
26. Sean $\mathbf{a}=\begin{pmatrix}1\\3\\2\end{pmatrix}$, $\mathbf{b}=\begin{pmatrix}-2\\4\\1\end{pmatrix}$ y $\mathbf{c}=\begin{pmatrix}0\\1\\4\end{pmatrix}$. Encuentre un vector $\mathbf{v}$ tal que $2\mathbf{a}-\mathbf{b}+3\mathbf{v}=4\mathbf{c}$.
27. Con $\mathbf{a}$, $\mathbf{b}$ y $\mathbf{c}$ como en el Problema 26, encuentre un vector $\mathbf{w}$ tal que $\mathbf{a}-\mathbf{b}+\mathbf{c}-\mathbf{w}=\mathbf{0}$.

1.4 Matrices

Matriz Una *matriz** A de $m \times n$ es un arreglo rectangular de mn números distribuidos en un orden de m renglones y n columnas:†

$$A=\begin{pmatrix}a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2n}\\ \vdots & \vdots & & \vdots & & \vdots\\ a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{in}\\ \vdots & \vdots & & \vdots & & \vdots\\ a_{m1} & a_{m2} & \cdots & a_{mj} & \cdots & a_{mn}\end{pmatrix} \tag{1}$$

* *Nota Histórica:* El término "matriz" fue usado por primera vez en 1850 por el matemático británico James Joseph Sylvester (1814-1897) para distinguir las matrices de los determinantes (que discutiremos en el Capítulo 2). De hecho, el término "matriz" quería significar "madre de los determinantes".

† Como en el caso de los vectores supondremos, si no se dice otra cosa, que los números en la matriz son reales.

El número a_{ij}, que aparece en el renglón i-ésimo y en la columna j-ésima de A, se conoce como la ij-ésima *componente* de A. Por conveniencia, la matriz A se escribe a veces $A = (a_{ij})$. Comúnmente las matrices se denotan por letras mayúsculas.

Si A es una matriz de $m\times n$ con $m=n$ se dice que A es una *matriz cuadrada*. Una matriz de $m\times n$ con todas sus componentes iguales a cero es una *matriz cero* de $m\times n$.

Se dice que una matriz de $m\times n$ tiene *tamaño* $m\times n$. Dos matrices $A=(a_{ij})$ y $B=(b_{ij})$ *son iguales* si (*i*) tienen el mismo tamaño y (*ii*) sus componentes correspondientes son iguales.

Ejemplo 1 A continuación se presentan cinco matrices de diferentes tamaños:

i. $A=\begin{pmatrix}1 & 3\\4 & 2\end{pmatrix}$, 2×2 (cuadrada) **ii.** $A=\begin{pmatrix}-1 & 3\\4 & 0\\1 & -2\end{pmatrix}$, 3×2

iii. $\begin{pmatrix}-1 & 4 & 1\\3 & 0 & 2\end{pmatrix}$, 2×3 **iv.** $\begin{pmatrix}1 & 6 & -2\\3 & 1 & 4\\2 & -6 & 5\end{pmatrix}$, 3×3 (cuadrada)

v. $\begin{pmatrix}0 & 0 & 0 & 0\\0 & 0 & 0 & 0\end{pmatrix}$, matriz cero de 2×4

Los vectores pueden ser vistos como casos especiales de matrices. Así, por ejemplo, el vector renglón n-dimensional $(a_1,a_2, \ldots ,a_n)$ es una matriz de $1\times n$, mientras que el vector columna n-dimensional $\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix}$ es una matriz de $n\times 1$.

Las matrices, como los vectores, surgen de un gran número de situaciones prácticas. Por ejemplo, vimos en la Sección 1.3 cómo el vector $\begin{pmatrix}10\\30\\15\\60\end{pmatrix}$ puede representar cantidades ordenadas de cuatro productos diferentes usados por un fabricante. Supongamos que hay cinco plantas distintas. Entonces la matriz de 4×5

$$Q=\begin{pmatrix}10 & 20 & 15 & 16 & 25\\30 & 10 & 20 & 25 & 22\\15 & 22 & 18 & 20 & 13\\60 & 40 & 50 & 35 & 45\end{pmatrix}$$

puede representar las órdenes de los cuatro productos en las cinco plantas. Podemos ver, por ejemplo, que la planta 4 ordena 25 unidades del segundo producto mientras que la planta 2 ordena 40 unidades del cuarto producto.

Las matrices, como los vectores, pueden sumarse y ser multiplicadas por escalares.*

Definición 1 **Suma de matrices.** Sean $A = (a_{ij})$ y $B = (b_{ij})$ dos matrices de $m \times n$. La suma de A y B es la matriz $A + B$ de $m \times n$ dada por

$$A+B=(a_{ij}+b_{ij})=\begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2n}+b_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \cdots & a_{mn}+b_{mn} \end{pmatrix} \quad (2)$$

Esto es, $A + B$ es la matriz de $m \times n$ obtenida al sumar las componentes correspondientes de A y B.

Advertencia. La suma de dos matrices está definida solamente cuando ambas matrices tienen el mismo tamaño. Así, por ejemplo, no es posible sumar entre las matrices $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$ y $\begin{pmatrix} -1 & 0 \\ 2 & -5 \\ 4 & 7 \end{pmatrix}$.

Ejemplo 2

$$\begin{pmatrix} 2 & 4 & -6 & 7 \\ 1 & 3 & 2 & 1 \\ -4 & 3 & -5 & 5 \end{pmatrix}+\begin{pmatrix} 0 & 1 & 6 & -2 \\ 2 & 3 & 4 & 3 \\ -2 & 1 & 4 & 4 \end{pmatrix}=\begin{pmatrix} 2 & 5 & 0 & 5 \\ 3 & 6 & 6 & 4 \\ -6 & 4 & -1 & 9 \end{pmatrix}$$

Definición 2 **Multiplicación de una matriz por un escalar.** Si $A = (a_{ij})$ es una matriz de $m \times n$ y si α es un escalar, entonces la matriz αA de $m \times n$ está dada por



Arthur Cayley (Biblioteca del Congreso, Estados Unidos)

* El álgebra de matrices, esto es, las reglas por medio de las cuales se pueden sumar y multiplicar matrices, fue desarrollada por el matemático inglés Arthur Cayley (1821-1895) en 1857. Con Cayley, las matrices surgieron en relación con transformaciones lineales del tipo

$$x' = ax + by,$$
$$y' = cx + dy,$$

en donde a, b, c, d son números reales, y que pueden considerarse como una representación (o mapeo) del punto (x, y) en el punto (x', y'). Claramente, la transformación se determina completamente por los cuatro coeficientes a, b, c, d, y entonces la transformación se simboliza por medio del arreglo cuadrado

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

al cual hemos llamado *matriz cuadrada*. Se analizarán las transformaciones lineales en el Capítulo 5.

$$\alpha A=(\alpha a_{ij})=\begin{pmatrix} \alpha a_{11} & \alpha a_{12} & \cdots & \alpha a_{1n} \\ \alpha a_{21} & \alpha a_{22} & \cdots & \alpha a_{2n} \\ \vdots & \vdots & & \vdots \\ \alpha a_{m1} & \alpha a_{m2} & \cdots & \alpha a_{mn} \end{pmatrix} \quad (3)$$

En otras palabras, $\alpha A = (\alpha a_{ij})$ es la matriz que se obtiene multiplicando por α cada componente de A.

Ejemplo 3 Sea $A=\begin{pmatrix} 1 & -3 & 4 & 2 \\ 3 & 1 & 4 & 6 \\ -2 & 3 & 5 & 7 \end{pmatrix}$. Entonce $2A=\begin{pmatrix} 2 & -6 & 8 & 4 \\ 6 & 2 & 8 & 12 \\ -4 & 6 & 10 & 14 \end{pmatrix}$,

$$-3A=\begin{pmatrix} -3 & 9 & -12 & -6 \\ -9 & -3 & -12 & -18 \\ 6 & -9 & -15 & -21 \end{pmatrix} \quad \text{y} \quad 0A=\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Ejemplo 4 Sea $A=\begin{pmatrix} 1 & 2 & 4 \\ -7 & 3 & -2 \end{pmatrix}$ y $B=\begin{pmatrix} 4 & 0 & 5 \\ 1 & -3 & 6 \end{pmatrix}$. Calcule $-2A+3B$.

Solución

$$-2A+3B=(-2)\begin{pmatrix} 1 & 2 & 4 \\ -7 & 3 & -2 \end{pmatrix}+(3)\begin{pmatrix} 4 & 0 & 5 \\ 1 & -3 & 6 \end{pmatrix}=$$
$$\begin{pmatrix} -2 & -4 & -8 \\ 14 & -6 & 4 \end{pmatrix}+\begin{pmatrix} 12 & 0 & 15 \\ 3 & -9 & 18 \end{pmatrix}=\begin{pmatrix} 10 & -4 & 7 \\ 17 & -15 & 22 \end{pmatrix}$$

El siguiente teorema es similar al Teorema 1.3.1. Su demostración queda como ejercicio (vea los Problemas del 21 al 24).

Teorema 1 Sean A, B y C matrices de $m \times n$ y sea α un escalar. Entonces:

i. $A + 0 = A$
ii. $0A = 0$ (Nótese que el cero de la izquierda es el número cero y el cero de la derecha es la matriz nula.)
iii. $A + B = B + A$ **(ley conmutativa para la suma de matrices)**
iv. $(A + B) + C = A + (B + C)$ **(ley asociativa para la suma de matrices)**
v. $\alpha(A + B) = \alpha A + \alpha B$ **(ley distributiva para la multiplicación por escalares)**
vi. $1A = A$

Nota. El cero en la parte (i) del teorema es la matriz cero de $m \times n$. En la parte (ii) el cero de la izquierda es un escalar, mientras que el cero de la derecha es la matriz cero de $m \times n$.

Ejemplo 5 Para ilustrar la ley asociativa notemos que

$$\left[\begin{pmatrix} 1 & 4 & -2 \\ 3 & -1 & 0 \end{pmatrix} + \begin{pmatrix} 2 & -2 & 3 \\ 1 & -1 & 5 \end{pmatrix}\right] + \begin{pmatrix} 3 & -1 & 2 \\ 0 & 1 & 4 \end{pmatrix} = \begin{pmatrix} 3 & 2 & 1 \\ 4 & -2 & 5 \end{pmatrix}$$

$$+ \begin{pmatrix} 3 & -1 & 2 \\ 0 & 1 & 4 \end{pmatrix} = \begin{pmatrix} 6 & 1 & 3 \\ 4 & -1 & 9 \end{pmatrix}$$

De igual manera,

$$\begin{pmatrix} 1 & 4 & -2 \\ 3 & -1 & 0 \end{pmatrix} + \left[\begin{pmatrix} 2 & -2 & 3 \\ 1 & -1 & 5 \end{pmatrix} + \begin{pmatrix} 3 & -1 & 2 \\ 0 & 1 & 4 \end{pmatrix}\right] = \begin{pmatrix} 1 & 4 & -2 \\ 3 & -1 & 0 \end{pmatrix}$$

$$+ \begin{pmatrix} 5 & -3 & 5 \\ 1 & 0 & 9 \end{pmatrix} = \begin{pmatrix} 6 & 1 & 3 \\ 4 & -1 & 9 \end{pmatrix}$$

Como con los vectores, la ley asociativa para la suma de matrices permite definir la suma de tres o más matrices.

Problemas 1.4

En los Problemas del 1 *al* 12 *efectúe las operaciones indicadas con*

$$A = \begin{pmatrix} 1 & 3 \\ 2 & 5 \\ -1 & 2 \end{pmatrix}, \; B = \begin{pmatrix} -2 & 0 \\ 1 & 4 \\ -7 & 5 \end{pmatrix} \text{ y } C = \begin{pmatrix} -1 & 1 \\ 4 & 6 \\ -7 & 3 \end{pmatrix}.$$

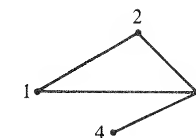
1. $3A$
2. $A+B$
3. $A-C$
4. $2C-5A$
5. $0B$ (0 es el escalar cero)
6. $-7A+3B$
7. $A+B+C$
8. $C-A-B$
9. $2A-3B+4C$
10. $7C-B+2A$
11. Encuentre una matriz D de manera que $2A+B-D$ sea la matriz cero de 3×2.
12. Obtenga una matriz E de manera que $A + 2B - 3C + E$ sea la matriz cero de 3×2.

En los Problemas del 13 *al* 20 *efectúe las operaciones indicadas con* $A =$

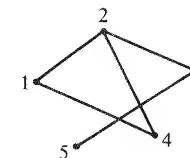
$$\begin{pmatrix} 1 & -1 & 2 \\ 3 & 4 & 5 \\ 0 & 1 & -1 \end{pmatrix}, \; B = \begin{pmatrix} 0 & 2 & 1 \\ 3 & 0 & 5 \\ 7 & -6 & 0 \end{pmatrix} \text{ y } C = \begin{pmatrix} 0 & 0 & 2 \\ 3 & 1 & 0 \\ 0 & -2 & 4 \end{pmatrix}.$$

13. $A-2B$
14. $3A-C$
15. $A+B+C$
16. $2A-B+2C$
17. $C-A-B$
18. $4C-2B+3A$

19. Halle un matriz D de manera que $A + B + C + D$ sea la matriz cero de 3×3.
20. Encuentre una matriz E de manera que $3C-2B+8A-4E$ sea la matriz cero de 3×3.
21. Sea $A = (a_{ij})$ una matriz de $m \times n$ y sea $\bar{0}$ la matriz cero de $m \times n$. Usando las Definiciones 1 y 2, muestre que $0A = \bar{0}$ y $\bar{0}+A = A$. Igualmente, muestre que $1A = A$.
22. Sean $A = (a_{ij})$ y $B = (b_{ij})$ matrices de $m \times n$. Calcule $A+B$ y $B+A$ y muestre que son iguales.
23. Si α es un escalar y A y B son como en el Problema 22, calcule $\alpha(A + B)$ y $\alpha A + \alpha B$ y muestre que son iguales.
24. Si $A = (a_{ij})$, $B = (b_{ij})$ y $C = (c_{ij})$ son matrices de $m \times n$, calcule $(A + B) + C$, y $A + (B + C)$ y muestre que son iguales.
25. Considere el "grafo" que une los cuatro puntos de la figura. Construya una matriz de 4×4 con la propiedad de que $a_{ij} = 0$ si el punto i no está conectado (por medio de una línea) con el punto j y $a_{ij} = 1$ si el punto i está conectado con el punto j.



26. Haga lo mismo (esta vez construya una matriz de 5×5) para el siguiente grafo.



1.5 Productos de vectores y matrices

En esta sección veremos cómo se pueden multiplicar dos matrices entre sí. Obviamente, podríamos definir el producto de dos matrices de $m \times n$, $A = (a_{ij})$ y $B = (b_{ij})$ como la matriz $m \times n$ cuya ij-ésima componente es $a_{ij}b_{ij}$. Sin embargo, debido a las importantes aplicaciones de las matrices, se necesita otra clase de producto. Comenzaremos definiendo el producto escalar de dos vectores.

Definición 1 ***Producto escalar de dos vectores.*** Sean $\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$ y $\mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$ dos n-vectores.

Entonces el *producto escalar* de $\mathbf{a}$ y $\mathbf{b}$, representado por $\mathbf{a} \cdot \mathbf{b}$, está dado por

$$\mathbf{a} \cdot \mathbf{b} = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n \qquad (1)$$

Por la notación en (1), el producto escalar de dos vectores frecuentemente es conocido como *producto punto* de los vectores. Notemos que el producto escalar de dos n-vectores es un escalar (esto es, un número).

Advertencia. Cuando hacemos el producto escalar de **a** y **b** necesitamos que **a** y **b** tengan el mismo número de componentes.

Frecuentemente efectuaremos el producto escalar de un vector renglón y un vector columna. En este caso tenemos

$$(a_1, a_2, \ldots, a_n) \cdot \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = a_1b_1 + a_2b_2 + \cdots + a_nb_n \qquad (2)$$

Ejemplo 1 Sean $\mathbf{a} = \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix}$ y $\mathbf{b} = \begin{pmatrix} 3 \\ -2 \\ 4 \end{pmatrix}$. Calcule $\mathbf{a} \cdot \mathbf{b}$.

Solución $\mathbf{a} \cdot \mathbf{b} = (1)(3) + (-2)(-2) + (3)(4) = 3 + 4 + 12 = 19$

Ejemplo 2 Sean $\mathbf{a} = (2, -3, 4, -6)$ y $\mathbf{b} = \begin{pmatrix} 1 \\ 2 \\ 0 \\ 3 \end{pmatrix}$. Calcule $\mathbf{a} \cdot \mathbf{b}$.

Solución Aquí $\mathbf{a} \cdot \mathbf{b} = (2)(1) + (-3)(2) + (4)(0) + (-6)(3) = 2 - 6 + 0 - 18 = -22$.

Ejemplo 3 Suponga que un fabricante produce cuatro artículos. La demanda para los artículos está dada por el vector de demanda $\mathbf{d} = (30, 20, 40, 10)$. Los precios unitarios para los artículos están dados por el vector de precios $\mathbf{p} = (\$20, \$15, \$18, \$40)$. Si satisface su demanda, ¿cuánto dinero recibirá el fabricante?

Solución La demanda del primer artículo es de 30 y el fabricante recibe \$20 por cada unidad vendida del primer artículo. Por lo tanto, recibe (30)(20) = \$600 por la venta del primer artículo. Continuando este razonamiento vemos que el total de dinero recibido será $\mathbf{d} \cdot \mathbf{p}$. Así, sus entradas son $\mathbf{d} \cdot \mathbf{p} = (30)(20) + (20)(15) + (40)(18) + (10)(40) = 600 + 300 + 720 + 400 = \$2020$.

El siguiente resultado se sigue directamente de la definición de producto escalar (Problema 22).

Teorema 1 Sean **a**, **b** y **c** n-vectores y sean α y β escalares. Entonces:

i. $\mathbf{a} \cdot \mathbf{0} = 0$
ii. $\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}$ **(ley conmutativa para el producto escalar)**
iii. $\mathbf{a} \cdot (\mathbf{b} + \mathbf{c}) = \mathbf{a} \cdot \mathbf{b} + \mathbf{a} \cdot \mathbf{c}$ **(ley distributiva para el producto escalar)**
iv. $(\alpha\mathbf{a}) \cdot \mathbf{b} = \alpha(\mathbf{a} \cdot \mathbf{b})$

Notemos que *no* hay ley asociativa para el producto escalar. La expresión $(\mathbf{a} \cdot \mathbf{b}) \cdot \mathbf{c} = \mathbf{a} \cdot (\mathbf{b} \cdot \mathbf{c})$ no tiene sentido ya que ninguno de los lados de la ecuación está definido. Para el lado izquierdo, esto se deduce del hecho de que $\mathbf{a} \cdot \mathbf{b}$ es un escalar y el producto escalar del escalar $\mathbf{a} \cdot \mathbf{b}$ y el vector **c** no está definido.

Definición 2 ***Producto de dos matrices.*** Sea $A = (a_{ij})$ una matriz de $m \times n$ cuyo i-ésimo renglón denotamos por $\mathbf{a}_i$. Sea $B = (b_{ij})$ una matriz de $n \times p$ cuya j-ésima columna denotamos por $\mathbf{b}_j$. Entonces el producto de A y B es una matriz $C = (c_{ij})$ de $m \times p$, donde

$$c_{ij} = \mathbf{a}_i \cdot \mathbf{b}_j \qquad (3)$$

Esto es, el ij-ésimo elemento de AB es el producto escalar del i-ésimo renglón de A ($\mathbf{a}_i$) y la j-ésima columna de B ($\mathbf{b}_j$). Si desarrollamos esto obtenemos

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} \qquad (4)$$

Advertencia. Dos matrices pueden multiplicarse sólo si el número de columnas de la primera es igual al número de renglones de la segunda. De otra forma, los vectores $\mathbf{a}_i$ y $\mathbf{b}_j$ tendrían diferente número de componentes y el producto escalar de la Ecuación (3) no estaría definido

---

Ejemplo 4 Si $A = \begin{pmatrix} 1 & 3 \\ -2 & 4 \end{pmatrix}$ y $B = \begin{pmatrix} 3 & -2 \\ 5 & 6 \end{pmatrix}$, calcule AB y BA.

Solución Sea $C = (c_{ij}) = AB$. Entonces $c_{11} = \mathbf{a}_1 \cdot \mathbf{b}_1 = (1 \quad 3) \cdot \begin{pmatrix} 3 \\ 5 \end{pmatrix} = 3 + 15 = 18$; $c_{12} = \mathbf{a}_1 \cdot \mathbf{b}_2 = (1 \quad 3) \cdot \begin{pmatrix} -2 \\ 6 \end{pmatrix} = -2 + 18 = 16$; $c_{21} = (-2 \quad 4) \cdot \begin{pmatrix} 3 \\ 5 \end{pmatrix} = -6 + 20 = 14$;

y $c_{22}=(-2 \quad 4)\cdot\begin{pmatrix}-2\\6\end{pmatrix}=4+24=28$. Así $C=AB=\begin{pmatrix}18 & 16\\14 & 28\end{pmatrix}$. Igualmente, dejando a un lado los pasos intermedios, vemos que

$$C'=BA=\begin{pmatrix}3 & -2\\5 & 6\end{pmatrix}\begin{pmatrix}1 & 3\\-2 & 4\end{pmatrix}=\begin{pmatrix}3+4 & 9-8\\5-12 & 15+24\end{pmatrix}=\begin{pmatrix}7 & 1\\-7 & 39\end{pmatrix}$$

Observación. El Ejemplo 4 ilustra un hecho importante: *el producto de matrices, en general, no es conmutativo*. Esto es, en general, $AB \neq BA$. En ocasiones sucede que $AB = BA$, pero esto será la excepción, no la regla. De hecho, como lo ilustra el siguiente ejemplo, puede ocurrir que AB esté definido mientras que BA no lo esté. Así, debemos ser cuidadosos en el *orden* en el cual multiplicamos dos matrices.

Ejemplo 5 Sean $A=\begin{pmatrix}2 & 0 & -3\\4 & 1 & 5\end{pmatrix}$ y $B=\begin{pmatrix}7 & -1 & 4 & 7\\2 & 5 & 0 & -4\\-3 & 1 & 2 & 3\end{pmatrix}$. Calcule AB.

Solución Primero notemos que A es una matriz de 2×3 y B es una matriz de 3×4. Por tanto el número de columnas de A es igual al número de renglones de B. El producto AB, por consiguiente está definido y es una matriz de 2×4. Sea $AB = C = (c_{ij})$. Entonces

$$c_{11}=(2 \quad 0 \quad -3)\cdot\begin{pmatrix}7\\2\\-3\end{pmatrix}=23 \qquad c_{12}=(2 \quad 0 \quad -3)\cdot\begin{pmatrix}-1\\5\\1\end{pmatrix}=-5$$

$$c_{13}=(2 \quad 0 \quad -3)\cdot\begin{pmatrix}4\\0\\2\end{pmatrix}=2 \qquad c_{14}=(2 \quad 0 \quad -3)\cdot\begin{pmatrix}7\\-4\\3\end{pmatrix}=5$$

$$c_{21}=(4 \quad 1 \quad 5)\cdot\begin{pmatrix}7\\2\\-3\end{pmatrix}=15 \qquad c_{22}=(4 \quad 1 \quad 5)\cdot\begin{pmatrix}-1\\5\\1\end{pmatrix}=6$$

$$c_{23}=(4 \quad 1 \quad 5)\cdot\begin{pmatrix}4\\0\\2\end{pmatrix}=26 \qquad c_{24}=(4 \quad 1 \quad 5)\cdot\begin{pmatrix}7\\-4\\3\end{pmatrix}=39$$

Por tanto $AB=\begin{pmatrix}23 & -5 & 2 & 5\\15 & 6 & 26 & 39\end{pmatrix}$. Esto termina el problema. Observe que el producto BA *no* está definido pues el número de columnas de B (cuatro) no es igual al número de renglones de A (dos).

Ejemplo 6 ***Contacto directo e indirecto con una enfermedad contagiosa.*** En este ejemplo se muestra como la multiplicación de matrices puede ser usada para modelar la propagación de una enfermedad contagiosa. Supóngase que cuatro individuos han contraído el padecimiento. Este grupo tiene contacto con seis personas de un segundo grupo. Podemos representar estos contactos, llamados *contactos directos*, por una matriz 4×6. Un ejemplo de una tal matriz lo damos a continuación:

Matriz de contactos directos: Grupos primero y segundo

$$A=\begin{pmatrix}0 & 1 & 0 & 0 & 1 & 0\\1 & 0 & 0 & 1 & 0 & 1\\0 & 0 & 0 & 1 & 1 & 0\\1 & 0 & 0 & 0 & 0 & 1\end{pmatrix}$$

Aquí se tiene que $a_{ij} = 1$ si la i-ésima persona del primer grupo ha hecho contacto con la j-ésima persona del segundo. Por ejemplo, el 1 en la posición 2, 4 significa que la segunda persona del primer grupo (el infectado) ha hecho contacto con la cuarta persona del segundo grupo. Ahora supóngase que un tercer grupo de cinco personas ha tenido contacto directo con individuos del segundo. Esto también puede representarse por una matriz.

Matriz de contactos directos: Grupos segundo y tercero,

$$B=\begin{pmatrix}0 & 0 & 1 & 0 & 1\\0 & 0 & 0 & 1 & 0\\0 & 1 & 0 & 0 & 0\\1 & 0 & 0 & 0 & 1\\0 & 0 & 0 & 1 & 0\\0 & 0 & 1 & 0 & 0\end{pmatrix}$$

Nótese que $b_{64} = 0$, lo cual significa que la sexta persona del segundo grupo no ha tenido contacto con la cuarta persona del tercero.

Los contactos *indirectos* —o *de segundo orden*— entre los individuos en el primero y tercer grupos se representan por la matriz 4×5 $C = AB$. Para captar esto, basta observar que para que alguien del grupo 3 haya sido contagiado, debió haber tenido contacto con alguien del grupo 2, quien a su vez hubo de ser contagiado por alguien del primer grupo. Por ejemplo, como $a_{24} = 1$ y $b_{45} = 1$, vemos que, indirectamente, la quinta persona del grupo 3 tuvo contacto (por medio de la cuarta persona del grupo 2) con la segunda persona del grupo 1. El número total de contactos indirectos entre la segunda persona del grupo 1 y la quinta del grupo 3 está dado por

$$\begin{aligned}c_{25} &= a_{21}b_{15} + a_{22}b_{25} + a_{23}b_{35} + a_{24}b_{45} + a_{25}b_{55} + a_{26}b_{65}\\ &= 1\cdot 1 + 0\cdot 0 + 0\cdot 0 + 1\cdot 1 + 0\cdot 0 + 1\cdot 0 = 2\end{aligned}$$

Ahora se calculará.

Matriz de contactos indirectos: Grupos primero y tercero,

$$C = AB = \begin{pmatrix} 0 & 0 & 0 & 2 & 0 \\ 1 & 0 & 2 & 0 & 2 \\ 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 2 & 0 & 1 \end{pmatrix}$$

Observamos que sólo la segunda persona en el grupo 3 no tiene contactos indirectos con enfermos. La quinta persona en este grupo tiene $2 + 1 + 1 = 4$ contactos indirectos.

Hemos visto que la ley conmutativa no se aplica al producto de matrices. El siguiente teorema muestra que la ley asociativa si se cumple.

Teorema 2 **Ley asociativa para multiplicación de matrices.** Sean $A = (a_{ij})$ una matriz de $n \times m$, $B = (b_{ij})$ una matriz de $m \times p$ y $C = (c_{ij})$ una matriz de $p \times q$. Entonces la *ley asociativa*

$$A(BC) = (AB)C \qquad (5)$$

es válida y ABC, definida por cualquier lado de (5) es una matriz de $n \times q$.

La demostración de este teorema no es difícil pero es algo tediosa. Es mejor darla usando notación de sumatoria. (Si esto no le es familiar, vea los Problemas del 60 al 80.) Por esta razón diferiremos la demostración hasta el final de la sección.

Ejemplo 7 Verifique la ley asociativa para $A = \begin{pmatrix} 1 & -3 \\ 0 & 2 \end{pmatrix}$, $B = \begin{pmatrix} 2 & -1 & 4 \\ 3 & 1 & 5 \end{pmatrix}$, y $C = \begin{pmatrix} 0 & -2 & 1 \\ 4 & 3 & 2 \\ -5 & 0 & 6 \end{pmatrix}$.

Solución Primero notemos que A es de 2×2, B es de 2×3 y C es de 3×3. Así, todos los productos empleados en el enunciado de la ley asociativa 1 están definidos y el producto resultante será una matriz de 2×3. Entonces calculamos

$$AB = \begin{pmatrix} 1 & -3 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} 2 & -1 & 4 \\ 3 & 1 & 5 \end{pmatrix} = \begin{pmatrix} -7 & -4 & -11 \\ 6 & 2 & 10 \end{pmatrix}$$

$$(AB)C = \begin{pmatrix} -7 & -4 & -11 \\ 6 & 2 & 10 \end{pmatrix}\begin{pmatrix} 0 & -2 & 1 \\ 4 & 3 & 2 \\ -5 & 0 & 6 \end{pmatrix} = \begin{pmatrix} 39 & 2 & -81 \\ -42 & -6 & 70 \end{pmatrix}$$

Igualmente,

$$BC = \begin{pmatrix} 2 & -1 & 4 \\ 3 & 1 & 5 \end{pmatrix}\begin{pmatrix} 0 & -2 & 1 \\ 4 & 3 & 2 \\ -5 & 0 & 6 \end{pmatrix} = \begin{pmatrix} -24 & -7 & 24 \\ -21 & -3 & 35 \end{pmatrix}$$

$$A(BC) = \begin{pmatrix} 1 & -3 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} -24 & -7 & 24 \\ -21 & -3 & 35 \end{pmatrix} = \begin{pmatrix} 39 & 2 & -81 \\ -42 & -6 & 70 \end{pmatrix}$$

Así, $(AB)C = A(BC)$.

De aquí en adelante escribiremos el producto de tres matrices simplemente como ABC. Podemos hacer esto porque $(AB)C = A(BC)$; así obtenemos el mismo resultado sin importar cómo se efectúe la multiplicación (con tal de que no conmutemos ninguna de las matrices).

La ley asociativa se puede extender a productos más grandes. Por ejemplo, si AB, BC y CD están definidos, entonces

$$ABCD = A(B(CD)) = ((AB)C)D = A(BC)D = (AB)(CD) \qquad (6)$$

Hay dos leyes distributivas para la multiplicación de matrices.

Teorema 3 **Leyes distributivas para la multiplicación de matrices.** Si todas las sumas y los productos siguientes están definidos, entonces

$$A(B + C) = AB + AC \qquad (7)$$

y

$$(A + B)C = AC + BC \qquad (8)$$

Demostración de los Teoremas 2 y 3 **Ley asociativa.** Como A es de $n \times m$ y B es de $m \times p$, AB es de $n \times p$. Así, $(AB)C = (n \times p) \times (p \times q)$ es una matriz de $n \times q$. Igualmente BC es de $m \times q$, y $A(BC)$ es de $n \times q$, y así $(AB)C$ y $A(BC)$ son ambas del mismo tamaño. Debemos mostrar que la ij-ésima componente de $(AB)C$ es igual a la ij-ésima componente de $A(BC)$. Definimos $D = (d_{ij}) = AB$. Entonces $d_{ij} = \sum_{k=1}^{m} a_{ik}b_{kj}$. La ij-ésima componente de $(AB)C = DC$ es $\sum_{l=1}^{p} d_{il}c_{lj} = \sum_{l=1}^{p}\left(\sum_{k=1}^{m} a_{ik}b_{kl}\right)c_{lj} = \sum_{k=1}^{m}\sum_{l=1}^{p} a_{ik}b_{kl}c_{lj}$. Ahora definimos $E = (e_{ij}) = BC$. Entonces $e_{ij} = \sum_{l=1}^{p} b_{il}c_{lj}$ y la ij-ésima componente de $A(BC) = AE$ es $\sum_{k=1}^{m} a_{ik}e_{kj} = \sum_{k=1}^{m}\sum_{l=1}^{p} a_{ik}b_{kl}c_{lj}$. Pero ésta es la ij-ésima componente de $(AB)C$ y el teorema queda así demostrado. ■

Leyes distributivas. Probaremos la primera ley distributiva (Ecuación 7). La prueba de la segunda (Ecuación 8) es idéntica y por lo tanto la omitimos. Sea A de $n \times m$ y sean B y C de $m \times p$. Entonces la kj-ésima componente de $B+C$ es $b_{kj}+c_{kj}$ y la ij-ésima componente de $A(B+C)$ es $\sum_{k=1}^{m} a_{ik}(b_{kj}+c_{kj}) = \sum_{k=1}^{m} a_{ik}b_{kj} + \sum_{k=1}^{m} a_{ik}c_{kj} = ij$-ésima componente de AB más la ij-ésima componente de AC, lo que prueba la Ecuación (7). ■

Problemas 1.5

En los Problemas del 1 al 7 calcule el producto escalar de los dos vectores.

1. $\begin{pmatrix} 2 \\ 3 \\ -5 \end{pmatrix}; \begin{pmatrix} 3 \\ 0 \\ 4 \end{pmatrix}$

2. $(1, 2, -1, 0);\ (3, -7, 4, -2)$

3. $\begin{pmatrix} 5 \\ 7 \end{pmatrix}; \begin{pmatrix} 3 \\ -2 \end{pmatrix}$

4. $(8, 3, 1);\ (7, -4, 3)$

5. $(a, b);\ (c, d)$

6. $\begin{pmatrix} x \\ y \\ z \end{pmatrix}; \begin{pmatrix} y \\ z \\ x \end{pmatrix}$

7. $(-1, -3, 4, 5);\ (-1, -3, 4, 5)$

8. Sea $\mathbf{a}$ un n-vector. Muestre que $\mathbf{a} \cdot \mathbf{a} \geq 0$.

9. Encuentre condiciones en un vector $\mathbf{a}$ de forma que $\mathbf{a} \cdot \mathbf{a} = 0$.

En los Problemas del 10 al 14 efectúe el cálculo indicado con $\mathbf{a} = \begin{pmatrix} 1 \\ -2 \\ 4 \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} 0 \\ -3 \\ -7 \end{pmatrix}$ y $\mathbf{c} = \begin{pmatrix} 4 \\ -1 \\ 5 \end{pmatrix}$.

10. $(2\mathbf{a}) \cdot (3\mathbf{b})$

11. $\mathbf{a} \cdot (\mathbf{b}+\mathbf{c})$

12. $\mathbf{c} \cdot (\mathbf{a}-\mathbf{b})$

13. $(2\mathbf{b}) \cdot (3\mathbf{c}-5\mathbf{a})$

14. $(\mathbf{a}-\mathbf{c}) \cdot (3\mathbf{b}-4\mathbf{a})$

En los Problemas del 15 al 29 efectúe las operaciones indicadas.

15. $\begin{pmatrix} 2 & 3 \\ -1 & 2 \end{pmatrix}\begin{pmatrix} 4 & 1 \\ 0 & 6 \end{pmatrix}$

16. $\begin{pmatrix} 3 & -2 \\ 1 & 4 \end{pmatrix}\begin{pmatrix} -5 & 6 \\ 1 & 3 \end{pmatrix}$

17. $\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} -1 & 0 \\ 2 & 3 \end{pmatrix}$

18. $\begin{pmatrix} -5 & 6 \\ 1 & 3 \end{pmatrix}\begin{pmatrix} 3 & -2 \\ 1 & 4 \end{pmatrix}$

19. $\begin{pmatrix} -4 & 5 & 1 \\ 0 & 4 & 2 \end{pmatrix}\begin{pmatrix} 3 & -1 & 1 \\ 5 & 6 & 4 \\ 0 & 1 & 2 \end{pmatrix}$

20. $\begin{pmatrix} 7 & 1 & 4 \\ 2 & -3 & 5 \end{pmatrix}\begin{pmatrix} 1 & 6 \\ 0 & 4 \\ -2 & 3 \end{pmatrix}$

21. $\begin{pmatrix} 1 & 6 \\ 0 & 4 \\ -2 & 3 \end{pmatrix}\begin{pmatrix} 7 & 1 & 4 \\ 2 & -3 & 5 \end{pmatrix}$

22. $\begin{pmatrix} 1 & 4 & -2 \\ 3 & 0 & 4 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 2 & 3 \end{pmatrix}$

23. $\begin{pmatrix} 1 & 4 & 6 \\ -2 & 3 & 5 \\ 1 & 0 & 4 \end{pmatrix}\begin{pmatrix} 2 & -3 & 5 \\ 1 & 0 & 6 \\ 2 & 3 & 1 \end{pmatrix}$

24. $\begin{pmatrix} 2 & -3 & 5 \\ 1 & 0 & 6 \\ 2 & 3 & 1 \end{pmatrix}\begin{pmatrix} 1 & 4 & 6 \\ -2 & 3 & 5 \\ 1 & 0 & 4 \end{pmatrix}$

25. $(1 \quad 4 \quad 0 \quad 2)\begin{pmatrix} 3 & -6 \\ 2 & 4 \\ 1 & 0 \\ -2 & 3 \end{pmatrix}$

26. $\begin{pmatrix} 3 & 2 & 1 & -2 \\ -6 & 4 & 0 & 3 \end{pmatrix}\begin{pmatrix} 1 \\ 4 \\ 0 \\ 2 \end{pmatrix}$

27. $\begin{pmatrix} 3 & -2 & 1 \\ 4 & 0 & 6 \\ 5 & 1 & 9 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

28. $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 3 & -2 & 1 \\ 4 & 0 & 6 \\ 5 & 1 & 9 \end{pmatrix}$

29. $\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & j \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, en donde $a, b, c, d, e, f, g, h, j$ son números reales.

30. Encuentre una matriz $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ tal que $A\begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

★ 31. Sean a_{11}, a_{12}, a_{21} y a_{22} números reales dados tales que $a_{11}a_{22} - a_{12}a_{21} \neq 0$. Encuentre números b_{11}, b_{12}, b_{21} y b_{22} tales que $\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

32. Verifique la ley asociativa de la multiplicación para las matrices $A = \begin{pmatrix} 2 & -1 & 4 \\ 1 & 0 & 6 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 0 & 1 \\ 2 & -1 & 2 \\ 3 & -2 & 0 \end{pmatrix}$, y $C = \begin{pmatrix} 1 & 6 \\ -2 & 4 \\ 0 & 5 \end{pmatrix}$.

33. Como en el Ejemplo 6, supóngase que un grupo de individuos ha contraído una enfermedad infecciosa. Estas personas tienen contacto con un segundo grupo que a su vez tiene contacto con un tercero. Sea $A = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}$ la matriz que representa los contactos entre el grupo contagioso y los miembros del grupo 2, y sea

$$B = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix}$$

la matriz que representa los contactos entre los grupos 2 y 3. (a) ¿Cuántas personas hay en cada grupo? (b) Halle la matriz de contactos indirectos entre los grupos 1 y 3.

34. Responda a las preguntas del Problema 33 si $A = \begin{pmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \end{pmatrix}$ y

$$B = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Vectores ortogonales *Se dice que dos vectores* **a** *y* **b** *son* ortogonales *si* $\mathbf{a} \cdot \mathbf{b} = 0$. *En los Problemas del 35 al 39 determine qué pares de vectores son ortogonales.**

35. $\begin{pmatrix} 2 \\ -3 \end{pmatrix}; \begin{pmatrix} 3 \\ 2 \end{pmatrix}$

36. $\begin{pmatrix} 2 \\ -3 \end{pmatrix}; \begin{pmatrix} -3 \\ 2 \end{pmatrix}$

37. $\begin{pmatrix} 1 \\ 4 \\ -7 \end{pmatrix}; \begin{pmatrix} 2 \\ 3 \\ 2 \end{pmatrix}$

38. $(1, 0, 1, 0);\ (0, 1, 0, 1)$

39. $\begin{pmatrix} a \\ 0 \\ b \\ 0 \\ c \end{pmatrix}; \begin{pmatrix} 0 \\ d \\ 0 \\ e \\ 0 \end{pmatrix}$

40. Encuentre un número α de forma que $(1, -2, 3, 5)$ sea ortogonal a $(-4, \alpha, 6, 1)$.

41. Encuentre todos los números α y β de forma que los vectores $\begin{pmatrix} 1 \\ -\alpha \\ 2 \\ 3 \end{pmatrix}$ y $\begin{pmatrix} 4 \\ 5 \\ -2\beta \\ 7 \end{pmatrix}$ sean ortogonales.

42. Demuestre el Teorema 1 a partir de la definición de producto escalar.

43. Un fabricante de joyería tiene pedidos para dos anillos, tres pares de aretes, cinco fistoles y un collar. El fabricante estima que requiere 1 hora de trabajo el elaborar un anillo; $1\frac{1}{2}$ horas el hacer un par de aretes; $\frac{1}{2}$ hora el hacer un fistol, y 2 horas, la elaboración de un collar.
   (a) Exprese las órdenes de trabajo o pedidos como un vector renglón.
   (b) Exprese los tiempos de elaboración de los diversos productos como un vector columna.
   (c) Utilice el producto escalar para determinar el número total de horas que se requerirán para surtir los pedidos.

44. Una turista que regresa de un viaje por Europa se encuentra que posee el siguiente dinero: 1000 chelines austriacos, 20 libras esterlinas, 100 francos franceses, 5000 liras italianas y 50 marcos alemanes. En dólares estadunidenses, un chelín valía \$0.055; la libra, \$1.80; el franco, \$0.20; la lira, \$0.001, y el marco, \$0.40.
   (a) Exprese la cantidad de cada moneda por un vector renglón.
   (b) Exprese el valor de cada moneda en dólares por un vector columna.
   (c) Utilice el producto escalar para hallar a cuántos dólares equivalía el dinero de la turista.

45. Una compañía les paga a sus ejecutivos su sueldo y les concede participación en las acciones como una gratificación anual. El año pasado, el presidente recibió 80,000 unidades monetarias (u.m.) y 50 acciones, cada uno de los tres vicepresidentes recibió 45,000 u.m. y 20 acciones, y el tesorero, 40,000 u.m. y 10 acciones.
   (a) Exprese los pagos en dinero y en acciones por medio de una matriz de 2×3.
   (b) Exprese el número de ejecutivos de cada categoría por un vector columna.
   (c) Utilice la multiplicación de matrices para calcular la cantidad total de dinero y de acciones que erogó la compañía en el pago a sus funcionarios principales el último año.

* Los vectores ortogonales serán tratados ampliamente en los Capítulos 3 y 4.

46. Las ventas, utilidades brutas por unidad y los impuestos unitarios de una gran compañía se dan en la tabla siguiente.

	Producto					
Mes	Ventas del Producto I	II	III	Producto	Utilidades (en cientos de u.m.)	Impuestos (en cientos de u.m.)
Enero	4	2	20	I	3.5	1.5
Febrero	6	1	9	II	2.75	2
Marzo	5	3	12	III	1.5	0.6
Abril	8	2.5	20			

Encuentre una matriz que muestre el total de utilidades e impuestos en cada uno de los 4 meses.

47. Sea A una matriz cuadrada. Entonces A^2 se define simplemente como AA. Calcule $\begin{pmatrix} 2 & -1 \\ 4 & 6 \end{pmatrix}^2$.

48. Calcule A^2, en donde $A = \begin{pmatrix} 1 & -2 & 4 \\ 2 & 0 & 3 \\ 1 & 1 & 5 \end{pmatrix}$.

49. Determine A^3, en donde $A = \begin{pmatrix} -1 & 2 \\ 3 & 4 \end{pmatrix}$.

50. Evalúe A^2, A^3, A^4 y A^5, en donde

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

51. Calcule A^2, A^3, A^4 y A^5, en donde

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

52. Una matriz A de $n \times n$ tiene la propiedad de que su producto con cualquier matriz de $n \times n$ es la matriz cero. Pruebe que A es la matriz cero.

53. Una *matriz de probabilidad* es una matriz cuadrada con dos propiedades: (*i*) cada componente es no negativa (≥ 0) y (*ii*) la suma de los elementos de cada renglón es 1. Las siguientes son matrices de probabilidad:

$$P = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ 0 & 0 & 1 \end{pmatrix} \quad \text{y} \quad Q = \begin{pmatrix} \frac{1}{6} & \frac{1}{6} & \frac{2}{3} \\ 0 & 1 & 0 \\ \frac{1}{5} & \frac{1}{5} & \frac{3}{5} \end{pmatrix}$$

Muestre que PQ es una matriz de probabilidad.

★54. Sea P una matriz de probabilidad. Muestre que P^2 es una matriz de probabilidad.

★55. Sean P y Q matrices de probabiliad del mismo tamaño. Pruebe que PQ es una matriz de probabilidad.

★56. Demuestre la fórmula (6) usando la ley asociativa (Ecuación 5).

★57. Un torneo de tenis puede ser organizado de la siguiente forma. Cada uno de los n jugadores juega contra todos los demás, y los resultados son registrados en una matriz R de $n \times n$ como sigue:

$$R_{ij} = \begin{cases} 1 \text{ si el } i\text{-ésimo jugador vence al } j\text{-ésimo jugador} \\ 0 \text{ si el } i\text{-ésimo jugador pierde ante el } j\text{-ésimo jugador} \\ 0 \text{ si } i = j \end{cases}$$

Al i-ésimo jugador se le asigna el siguiente puntaje

$$S_i = \sum_{j=1}^{n} R_{ij} + \frac{1}{2}\sum_{j=1}^{n} (R^2)_{ij}\,^*$$

(a) En un torneo entre cuatro jugadores.

$$R = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{pmatrix}.$$

Clasifique a los jugadores de acuerdo a sus puntajes.

(b) Interprete el significado de la clasificación.

58. Sea O la matriz cero de $m \times n$ y sea A una matriz de $n \times p$. Muestre que $OA = O_1$, donde O_1 es la matriz cero de $m \times p$.

59. Verifique la ley distributiva (Ecuación 5) para las matrices

$$A = \begin{pmatrix} 1 & 2 & 4 \\ 3 & -1 & 0 \end{pmatrix} \quad B = \begin{pmatrix} 2 & 7 \\ -1 & 4 \\ 6 & 0 \end{pmatrix} \quad C = \begin{pmatrix} -1 & 2 \\ 3 & 7 \\ 4 & 1 \end{pmatrix}$$

Debido a que usaremos la notación Σ en varias partes de este libro, el lector debería familiarizarse con ella. Los siguientes problemas están diseñados para este propósito. En los Problemas del 60 al 67 evalúe las sumas dadas.

60. $\sum_{k=1}^{4} 2k$ 61. $\sum_{i=1}^{3} i^3$ 62. $\sum_{k=0}^{6} 1$ 63. $\sum_{k=1}^{8} 3^k$

64. $\sum_{i=2}^{5} \frac{1}{1+i}$ 65. $\sum_{j=5}^{7} \frac{2j+3}{j-2}$ 66. $\sum_{i=1}^{3}\sum_{j=1}^{4} ij$ 67. $\sum_{k=1}^{3}\sum_{j=2}^{4} k^2 j^3$

En los Problemas del 68 al 80 escriba cada suma usando la notación Σ.

68. $1+2+4+8+16$ 69. $1-3+9-27+81-243$

70. $\frac{2}{3}+\frac{3}{4}+\frac{4}{5}+\frac{5}{6}+\frac{6}{7}+\frac{7}{8}+\cdots+\frac{n}{n+1}$ 71. $1+2^{1/2}+3^{1/3}+4^{1/4}+5^{1/5}+\cdots+n^{1/n}$

72. $1+x^3+x^6+x^9+x^{12}+x^{15}+x^{18}+x^{21}$

73. $-1+\frac{1}{a}-\frac{1}{a^2}+\frac{1}{a^3}-\frac{1}{a^4}+\frac{1}{a^5}-\frac{1}{a^6}+\frac{1}{a^7}-\frac{1}{a^8}+\frac{1}{a^9}$

74. $1\cdot 3+3\cdot 5+5\cdot 7+7\cdot 9+9\cdot 11+11\cdot 13+13\cdot 15+15\cdot 17$

* $(R^2)_{ij}$ es la ij-ésima componente de la matriz R^2.

75. $2^2\cdot 4+3^2\cdot 6+4^2\cdot 8+5^2\cdot 10+6^2\cdot 12+7^2\cdot 14$

76. $a_{11}+a_{12}+a_{13}+a_{21}+a_{22}+a_{23}$

77. $a_{11}+a_{12}+a_{21}+a_{22}+a_{31}+a_{32}$

78. $a_{21}+a_{22}+a_{23}+a_{24}+a_{31}+a_{32}+a_{33}+a_{34}+a_{41}+a_{42}+a_{43}+a_{44}$

79. $a_{31}b_{12}+a_{32}b_{22}+a_{33}b_{32}+a_{34}b_{42}+a_{35}b_{52}$

80. $a_{21}b_{11}c_{15}+a_{21}b_{12}c_{25}+a_{21}b_{13}c_{35}+a_{21}b_{14}c_{45}$
$+a_{22}b_{21}c_{15}+a_{22}b_{22}c_{25}+a_{22}b_{23}c_{35}+a_{22}b_{24}c_{45}$
$+a_{23}b_{31}c_{15}+a_{23}b_{32}c_{25}+a_{23}b_{33}c_{35}+a_{23}b_{34}c_{45}$

1.6 m Ecuaciones en n incógnitas: Eliminaciones de Gauss-Jordan y gaussiana

En esta sección describimos un método para encontrar todas las soluciones (si existen) de un sistema de m ecuaciones lineales en n incógnitas. Al hacer esto veremos que, como en el caso de 2×2, tales sistemas tienen una solución, un número infinito de soluciones o ninguna solución. Antes de entrar al método en general veamos algunos ejemplos simples.

Ejemplo 1 Resuelva el sistema

$$\begin{aligned} 2x_1 + 4x_2 + 6x_3 &= 18 \\ 4x_1 + 5x_2 + 6x_3 &= 24 \\ 3x_1 + x_2 - 2x_3 &= 4 \end{aligned} \tag{1}$$

Solución Buscamos tres números x_1, x_2 y x_3 tales que las tres ecuaciones en (1) sean satisfechas. Nuestro método de solución consistirá en simplificar las ecuaciones como lo hicimos en la Sección 1.2 de forma que las soluciones sean fácilmente identificadas. Empezamos por dividir la primera ecuación entre 2. Esto da

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 9 \\ 4x_1 + 5x_2 + 6x_3 &= 24 \\ 3x_1 + x_2 - 2x_3 &= 4 \end{aligned} \tag{2}$$

Como vimos en la sección antes mencionada, sumar dos ecuaciones nos lleva a una tercera ecuación válida. Esta ecuación puede reemplazar en el sistema a cualquiera de las dos ecuaciones usadas para obtenerla. Empezamos a simplificar el sistema (2) multiplicando por −4 ambos lados de la primera ecuación en (2) y sumando esta nueva ecuación a la segunda. Esto nos da

$$\begin{aligned} -4x_1 - 8x_2 - 12x_3 &= -36 \\ 4x_1 + 5x_2 + 6x_3 &= 24 \\ \hline -3x_2 - 6x_3 &= -12 \end{aligned}$$

La ecuación $-3x_2 - 6x_3 = -12$ es nuestra segunda nueva ecuación y ahora el sistema es

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 9 \\ -3x_2 - 6x_3 &= -12 \\ 3x_1 + x_2 - 2x_3 &= 4 \end{aligned}$$

Ahora multiplicamos la primera ecuación por −3 y la sumamos a la tercera ecuación:

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 9 \\ -3x_2 - 6x_3 &= -12 \\ -5x_2 - 11x_3 &= -23 \end{aligned} \qquad (3)$$

Notemos que en el sistema (3) la variable x_1 ha sido eliminada de la segunda y tercera ecuaciones. Después dividimos la segunda ecuación entre −3:

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 9 \\ x_2 + 2x_3 &= 4 \\ -5x_2 - 11x_3 &= -23 \end{aligned}$$

Multiplicamos la segunda ecuación por −2 y la sumamos a la primera; luego se multiplica la segunda ecuación por 5 y se suma a la tercera:

$$\begin{aligned} x_1 \qquad - x_3 &= 1 \\ x_2 + 2x_3 &= 4 \\ - x_3 &= -3 \end{aligned}$$

Multiplicando la tercera ecuación por −1:

$$\begin{aligned} x_1 \qquad - x_3 &= 1 \\ x_2 + 2x_3 &= 4 \\ x_3 &= 3 \end{aligned}$$

Finalmente, sumamos la tercera ecuación a la primera y después se multiplica la tercera ecuación por −2 y se la suma a la segunda para obtener el siguiente sistema (el cual es equivalente al sistema (1)):

$$\begin{aligned} x_1 &= 4 \\ x_2 &= -2 \\ x_3 &= 3 \end{aligned}$$

Ésta es la solución única al sistema. La escribimos en la forma de vector (4, −2, 3). El método aquí usado se conoce como método de *eliminación de Gauss-Jordan*.*

* Llamado así en honor del gran matemático alemán Carl Friedrich Gauss (1777-1855) y del matemático francés Camille Jordan (1838-1922). Véase la reseña biográfica de Gauss.

Antes de continuar con otro ejemplo resumiremos lo hecho en éste:

i. Dividimos para hacer igual a 1 el coeficiente de x_1 en la primera ecuación.
ii. "Eliminamos" los términos en x_1 de la segunda y tercera ecuaciones. Esto es, hicimos los coeficientes de estos términos iguales a cero multiplicando la primera ecuación por números apropiados y después sumándola a la segunda y tercera ecuaciones, respectivamente.
iii. Dividimos para hacer igual a 1 el coeficiente del término en x_2 en la segunda ecuación y después usamos ésta para eliminar los términos en x_2 de la primera y tercera ecuaciones.
iv. Dividimos para hacer igual a 1 el coeficiente del término en x_3 de la tercera ecuación y después usamos esta ecuación para eliminar los términos en x_3 de la primera y segunda ecuaciones.

Debemos enfatizar que en cada paso obtuvimos sistemas que son equivalentes. Esto es, cada sistema tiene el mismo conjunto de soluciones que el sistema anterior. Esto se deduce de las Propiedades A y B de la Sección 1.2.

Matriz de coeficientes

Antes de resolver otros sistemas de ecuaciones, introduciremos una notación que haga más fácil escribir cada paso en nuestro procedimiento. Una *matriz* es un arreglo rectangular de números. Los coeficientes de las variables x_1, x_2 y x_3 en el sistema (1) pueden ser escritos como los elementos de una matriz A, llamada *matriz de coeficientes* del sistema:

$$A = \begin{pmatrix} 2 & 4 & 6 \\ 4 & 5 & 6 \\ 3 & 1 & -2 \end{pmatrix} \qquad (4)$$

Defina los vectores **x** y **b** como

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \quad \text{y} \quad \mathbf{b} = \begin{pmatrix} 18 \\ 24 \\ 4 \end{pmatrix}$$

El sistema (1) puede escribirse en la forma

$$A\mathbf{x} = \mathbf{b}$$

Matriz aumentada

En cada paso del Ejemplo 1, se escriben muchas x. Esto no es necesario: Usando la notación matricial, el sistema (1) puede expresarse como la *matriz aumentada*

$$\left(\begin{array}{ccc|c} 2 & 4 & 6 & 18 \\ 4 & 5 & 6 & 24 \\ 3 & 1 & -2 & 4 \end{array}\right) \qquad (5)$$

Por ejemplo, la primera fila de la matriz aumentada (5) se lee $2x_1 + 4x_2 + 6x_3 = 18$. Nótese que cada fila o renglón de la matriz aumentada corresponde a una de las ecuaciones del sistema.

Ahora introduciremos alguna terminología. Hemos visto que multiplicando (o dividiendo) los dos lados de una ecuación por un número diferente de cero

Carl Friedrich Gauss, 1777-1855



Carl Friedrich Gauss (Biblioteca del Congreso, Estados Unidos)

El más grande matemático del siglo diecinueve, Carl Friedrich Gauss es considerado uno de los tres más grandes matemáticos de todos los tiempos —siendo los otros dos Arquímedes y Newton.

Gauss nació en Brunsvick, Alemania en 1777. Su padre, un trabajador pertinaz que fue excepcionalmente obstinado y nunca creyó en la educación formal, hizo todo lo que pudo para mantener a Gauss fuera de las escuelas. Afortunadamente para Carl (y para las matemáticas), su madre, pese a no contar tampoco con educación formal, alentó a su hijo en sus estudios y participó con orgullo de sus logros hasta los días de su muerte, a la edad de 97 años.

Gauss fue un niño prodigio. A la edad de 3 años, encontró un error en la contabilidad de su padre. Se cuenta una famosa historia de Carl, a la edad de 10 años, cuando pupilo de la escuela local de Brunsvick. El maestro era famoso por sus asignaciones de tareas para mantener a los niños ocupados. Un buen día les pidió ponerse a sumar los números del 1 al 100. Casi en forma instantánea, Carl colocó su pizarra cara abajo con las palabras "ahí está". Después, el maestro descubrió que Carl era el único que tenía la respuesta correcta, 5050. Gauss había notado que los números podían disponerse en 50 parejas, cada una con la suma 101 (1 + 100, 2 + 99, y así sucesivamente), y 50 × 101 = 5050. Más tarde en su vida, Gauss decía que antes de aprender a hablar, ya sabía sumar.

Cuando tenía 15 años, el duque de Brunsvick advirtió su talento, y se convirtió en su protector. El duque le ayudó a ingresar a la Universidad de Brunsvick en 1795, y tres años después, a la Universidad de Gotinga. Titubeante entre las disciplinas de filosofía y matemática, finalmente eligió las ciencias matemáticas después de dos extraordinarios descubrimientos. Primero, inventó el método de mínimos cuadrados, una década antes de que Legendre publicara el resultado. Segundo, un mes antes de su cumpleaños decimonono, resolvió un problema cuya solución se había buscado por más de dos mil años. Gauss mostró cómo construir, con sólo compás y regla, polígonos regulares cuyo número de lados no fuese un múltiplo de 2, 3 o 5. El 30 de marzo de 1796, día de su descubrimiento, comenzó un diario cuyas líneas iniciales describían la construcción de un polígono regular de 17 lados. El diario, que contiene 146 menciones de resultados en sólo 19 páginas, es uno de los documentos más importantes en la historia de las matemáticas.

Después de un periodo breve en Gotinga, Gauss pasó a la Universidad de Helmstaedt y en 1798, a la edad de 20 años, escribió su famosa disertación doctoral. En ella dio la primera demostración rigurosa del teorema fundamental del algebra: todo polinomio de grado n, posee, contando las multiplicidades, n raíces. Muchos matemáticos, incluyendo a Euler, Newton y Lagrange, habían intentado la demostración de este resultado.

Gauss realizó un gran número de descubrimientos en física y en matemáticas. Por ejemplo, en 1801 usó un nuevo procedimiento para calcular, con muy pocos datos, la órbita del asteroide Ceres. En 1833, inventó el telégrafo electromagnético con su colega Wilhelm Weber (1804-1891). No obstante su brillante trabajo en astronomía y electricidad, su desempeño en matemáticas es todavía más asombroso. Hizo contribuciones fundamentales al álgebra y la geometría. En 1811, descubrió un resultado que indujo a Cauchy a desarrollar la teoría de la variable compleja. Lo encontramos aquí (en este libro) en el método de eliminación llamado de Gauss-Jordan. Los estudiantes de análisis numérico estudian la cuadratura de Gauss —una técnica de integración numérica.

Gauss se convirtió en catedrático de matemáticas en Gotinga en 1807, y permaneció en ese puesto hasta su muerte en 1855. Aun después de su muerte, su espíritu matemático quedó para tormento de los matemáticos del siglo diecinueve. Frecuentemente ocurría que algún supuesto nuevo resultado importante ya había sido descubierto anteriormente por Gauss y podía ser encontrado en sus notas no publicadas.

En sus escritos de matemáticas, Gauss era un perfeccionista y es quizá el último matemático que sabía todo de su materia. Defendiendo el punto de vista de que una catedral que se construye no es tal sino hasta que la última parte del andamiaje se ha retirado, hacía cada una de sus obras completa, concisa y pulida. Utilizaba un sello que mostraba un árbol con algunos pocos frutos y el lema: *Pauca sed matura* (pocos pero maduros). Pero Gauss también creyó que las matemáticas debían reflejar el mundo real. A su muerte, fue honrado con una medalla conmemorativa en la que se inscribió "De Jorge V, Rey de Hánover, al Príncipe de los Matemáticos".

se obtiene una nueva ecuación válida. Más aun, si sumamos un múltiplo de una ecuación a otra ecuación del sistema obtenemos otra ecuación válida. Finalmente, si intercambiamos dos de las ecuaciones de un sistema se obtiene un sistema equivalente. Estas tres operaciones, cuando se le aplican a los renglones de la matriz aumentada de un sistema de ecuaciones, se denominan *operaciones elementales de renglón* (o *sobre renglones*).

Resumiendo, las tres operaciones elementales sobre renglones aplicables a la representación de la matriz aumentada de un sistema de ecuaciones son:

Operaciones elementales de renglón

i. Multiplicar (o dividir) un renglón por un número distinto de cero.
ii. Sumar un múltiplo de un renglón a otro renglón.
iii. Intercambiar dos renglones.

El proceso de efectuar operaciones elementales de renglón para simplificar una matriz aumentada se llama *reducción por renglones.*

Notación

i. $M_i(c)$ indica "multiplicar el i-ésimo renglón de una matriz por el número c".
ii. $A_{i,j}(c)$ indica "multiplicar el i-ésimo renglón por c y sumárselo al j-ésimo renglón".
iii. $P_{i,j}$ indica "intercambiar (permutar) los renglones i y j".
iv. $A \to B$ indica que las matrices aumentadas A y B son equivalentes; esto es, que los sistemas que representan tienen la misma solución.

En el Ejemplo 1 vimos que si usamos las operaciones elementales de renglón (i) y (ii) varias veces obtenemos un sistema en el cual las soluciones al sistema quedan dadas explícitamente. Ahora repetimos los pasos del Ejemplo 1, usando la notación anterior:

$$\left(\begin{array}{ccc|c} 2 & 4 & 6 & 18 \\ 4 & 5 & 6 & 24 \\ 3 & 1 & -2 & 4 \end{array}\right) \xrightarrow{M_1(\frac{1}{2})} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 9 \\ 4 & 5 & 6 & 24 \\ 3 & 1 & -2 & 4 \end{array}\right) \xrightarrow[A_{1,3}(-3)]{A_{1,2}(-4)} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 9 \\ 0 & -3 & -6 & -12 \\ 0 & -5 & -11 & -23 \end{array}\right)$$

$$\xrightarrow{M_2(-\frac{1}{3})} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 9 \\ 0 & 1 & 2 & 4 \\ 0 & -5 & -11 & -23 \end{array}\right) \xrightarrow[A_{2,3}(5)]{A_{2,1}(-2)} \left(\begin{array}{ccc|c} 1 & 0 & -1 & 1 \\ 0 & 1 & 2 & 4 \\ 0 & 0 & -1 & -3 \end{array}\right)$$

$$\xrightarrow{M_3(-1)} \left(\begin{array}{ccc|c} 1 & 0 & -1 & 1 \\ 0 & 1 & 2 & 4 \\ 0 & 0 & 1 & 3 \end{array}\right) \xrightarrow[A_{3,2}(-2)]{A_{3,1}(1)} \left(\begin{array}{ccc|c} 1 & 0 & 0 & 4 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 3 \end{array}\right)$$

De nuevo fácilmente "vemos" la solución $x_1 = 4$, $x_2 = -2$, $x_3 = 3$.

Ejemplo 2 Resolver el sistema

$$\begin{aligned} 2x_1 + 4x_2 + 6x_3 &= 18 \\ 4x_1 + 5x_2 + 6x_3 &= 24 \\ 2x_1 + 7x_2 + 12x_3 &= 30 \end{aligned}$$

Solución Observemos que este sistema puede escribirse como $A\mathbf{x} = \mathbf{b}$, en donde

$$A = \begin{pmatrix} 2 & 4 & 6 \\ 4 & 5 & 6 \\ 2 & 7 & 12 \end{pmatrix}, \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \quad \text{y} \quad \mathbf{b} = \begin{pmatrix} 18 \\ 24 \\ 30 \end{pmatrix}$$

Para resolverlo procedemos como en el Ejemplo 1, expresando primero el sistema como una matriz aumentada:

$$\left(\begin{array}{ccc|c} 2 & 4 & 6 & 18 \\ 4 & 5 & 6 & 24 \\ 2 & 7 & 12 & 30 \end{array}\right)$$

Se obtiene sucesivamente,

$$\xrightarrow{M_1(\frac{1}{2})} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 9 \\ 4 & 5 & 6 & 24 \\ 2 & 7 & 12 & 30 \end{array}\right) \xrightarrow[A_{1,3}(-2)]{A_{1,2}(-4)} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 9 \\ 0 & -3 & -6 & -12 \\ 0 & 3 & 6 & 12 \end{array}\right)$$

$$\xrightarrow{M_2(-\frac{1}{3})} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 9 \\ 0 & 1 & 2 & 4 \\ 0 & 3 & 6 & 12 \end{array}\right) \xrightarrow[A_{2,3}(-3)]{A_{2,1}(-2)} \left(\begin{array}{ccc|c} 1 & 0 & -1 & 1 \\ 0 & 1 & 2 & 4 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Esto equivale al sistema de ecuaciones

$$\begin{aligned} x_1 \qquad - x_3 &= 1 \\ x_2 + 2x_3 &= 4 \end{aligned}$$

Hasta aquí puede llegarse. Hay sólo dos ecuaciones en tres incógnitas x_1, x_2, x_3 y puede haber un número infinito de soluciones. Para ver esto, elíjase x_3. Entonces $x_2 = 4 - 2x_3$ y $x_1 = 1 + x_3$. Esto es una solución para cualquier número x_3. Escribimos estas soluciones en la forma vectorial $(1 + x_3, 4 - 2x_3, x_3)$. Por ejemplo, si $x_3 = 0$, se obtiene la solución (1, 4, 0). Para $x_3 = 10$, resulta la solución (11, −16, 10).

Ejemplo 3 Resuelva el sistema

$$\begin{aligned} 2x_1 + 4x_2 + 6x_3 &= 18 \\ 4x_1 + 5x_2 + 6x_3 &= 24 \\ 2x_1 + 7x_2 + 12x_3 &= 40 \end{aligned} \tag{6}$$

Solución Usaremos la forma de la matriz aumentada y procederemos exactamente como en el Ejemplo 2 para obtener, sucesivamente, los siguientes sistemas. (Notemos que en cada paso, usamos las operaciones (*i*) o (*ii*).)

$$\left(\begin{array}{ccc|c} 2 & 4 & 6 & 18 \\ 4 & 5 & 6 & 24 \\ 2 & 7 & 12 & 40 \end{array}\right) \xrightarrow{M_1(\frac{1}{2})} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 9 \\ 4 & 5 & 6 & 24 \\ 2 & 7 & 12 & 40 \end{array}\right)$$

$$\xrightarrow[A_{1,3}(-2)]{A_{1,2}(-4)} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 9 \\ 0 & -3 & -6 & -12 \\ 0 & 3 & 6 & 22 \end{array}\right) \xrightarrow{M_2(-\frac{1}{3})} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 9 \\ 0 & 1 & 2 & 4 \\ 0 & 3 & 6 & 22 \end{array}\right)$$

$$\xrightarrow[A_{2,3}(-3)]{A_{2,1}(-2)} \left(\begin{array}{ccc|c} 1 & 0 & -1 & 1 \\ 0 & 1 & 2 & 4 \\ 0 & 0 & 0 & 10 \end{array}\right) \xrightarrow{M_3(\frac{1}{10})} \left(\begin{array}{ccc|c} 1 & 0 & -1 & 1 \\ 0 & 1 & 2 & 4 \\ 0 & 0 & 0 & 1 \end{array}\right)$$

La última ecuación queda $0x_1 + 0x_2 + 0x_3 = 1$, lo que es imposible ya que $0 \neq 1$. Así, el sistema (6) *no* tiene solución.

Veamos otra vez estos tres ejemplos. En el Ejemplo 1 empezamos con la matriz

$$A_1 = \begin{pmatrix} 2 & 4 & 6 \\ 4 & 5 & 6 \\ 3 & 1 & -2 \end{pmatrix}$$

En el proceso de reducción por renglones A_1 fue "reducida" a la matriz

$$R_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

En el Ejemplo 2 empezamos con

$$A_2 = \begin{pmatrix} 2 & 4 & 6 \\ 4 & 5 & 6 \\ 2 & 7 & 12 \end{pmatrix}$$

y terminamos con

$$R_2 = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

En el Ejemplo 3 empezamos con

$$A_3 = \begin{pmatrix} 2 & 4 & 6 \\ 4 & 5 & 6 \\ 2 & 7 & 12 \end{pmatrix}$$

y, otra vez, terminamos con

$$R_3 = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

Las matrices R_1, R_2 y R_3 se denominan *formas escalonadas reducidas* de las matrices A_1, A_2 y A_3, respectivamente. En general, tenemos la siguiente definición.

Definición 1 **Forma escalonada reducida.** Una matriz está en *forma escalonada reducida* si se cumplen las cuatro condiciones siguientes:

- **i.** Todos los renglones que consisten únicamente de ceros (si existen) aparecen en la parte de abajo de la matriz.
- **ii.** El primer número distinto de cero (si empezamos por la izquierda) en cualquier renglón que no consista únicamente de ceros es 1.
- **iii.** Si dos renglones sucesivos no consisten únicamente de ceros, entonces el primer 1 en el renglón inferior está más a la derecha que el primer 1 del renglón superior.
- **iv.** Cualquier columna que contenga el primer 1 de un renglón tendrá ceros en los demás lugares.

Ejemplo 4 Las siguientes matrices están en forma escalonada reducida:

$$\textbf{i.}\ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \textbf{ii.}\ \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \textbf{iii.}\ \begin{pmatrix} 1 & 0 & 0 & 5 \\ 0 & 0 & 1 & 2 \end{pmatrix}$$

$$\textbf{iv.}\ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \textbf{v.}\ \begin{pmatrix} 1 & 0 & 2 & 5 \\ 0 & 1 & 3 & 6 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Definición 2 **Forma escalonada.** Una matriz está en *forma escalonada* si se cumplen las condiciones (*i*), (*ii*) y (*iii*) de la Definición 1.

Ejemplo 5 Las siguientes matrices están en forma escalonada:

$$\textbf{i.}\ \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 5 \\ 0 & 0 & 1 \end{pmatrix} \quad \textbf{ii.}\ \begin{pmatrix} 1 & -1 & 6 & 4 \\ 0 & 1 & 2 & -8 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\textbf{iii.}\ \begin{pmatrix} 1 & 0 & 2 & 5 \\ 0 & 0 & 1 & 2 \end{pmatrix} \quad \textbf{iv.}\ \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \quad \textbf{v.}\ \begin{pmatrix} 1 & 3 & 2 & 5 \\ 0 & 1 & 3 & 6 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Observación 1. La diferencia entre estas dos formas debe ser clara a partir de los ejemplos. En la forma escalonada todos los números que están abajo del primer 1 de un renglón son cero. En la forma escalonada reducida todos los números que están arriba y abajo del primer 1 de un renglón son cero. Así, la forma escalonada reducida es más exclusiva. Esto es, cualquier matriz en forma escalonada reducida está en forma escalonada pero no inversamente.

Observación 2. Siempre podemos reducir una matriz a la forma escalonada reducida o a la forma escalonada realizando operaciones elementales de renglón. Hemos visto esta reducción a la forma escalonada reducida en los Ejemplos 1, 2 y 3.

Como hemos visto en los Ejemplos 1, 2 y 3, existe una fuerte conexión entre la forma escalonada reducida de una matriz y la existencia de una única solución al sistema. En el Ejemplo 1, la forma escalonada reducida de la *matriz de coeficientes* (esto es, las primeras tres columnas de la matriz aumentada) tienen un 1 en cada renglón y existe una única solución. En los Ejemplos 2 y 3, la forma escalonada reducida de la matriz de coeficientes tiene un renglón de ceros y el sistema no tiene solución o tiene un número infinito de soluciones. Esto siempre resulta ser cierto en cualquier sistema con el mismo número de ecuaciones que de incógnitas. Antes de ir al caso general, discutiremos la utilidad de la forma escalonada de una matriz. Es posible resolver el sistema del Ejemplo 1 reduciendo la matriz de coeficientes a su forma escalonada.

Ejemplo 6 Resuelva el sistema del Ejemplo 1 reduciendo la matriz de coeficientes a su forma escalonada.

Solución Empezamos como antes:

$$\left(\begin{array}{ccc|c} 2 & 4 & 6 & 18 \\ 4 & 5 & 6 & 24 \\ 3 & 1 & -2 & 4 \end{array}\right) \xrightarrow{M_1(\frac{1}{2})} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 9 \\ 4 & 5 & 6 & 24 \\ 3 & 1 & -2 & 4 \end{array}\right) \xrightarrow[A_{1,3}(-3)]{A_{1,2}(-4)}$$

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 9 \\ 0 & -3 & -6 & -12 \\ 0 & -5 & -11 & -23 \end{array}\right) \xrightarrow{M_2(-\frac{1}{3})} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 9 \\ 0 & 1 & 2 & 4 \\ 0 & -5 & -11 & -23 \end{array}\right)$$

Hasta ahora, el proceso es el mismo que se hizo anteriormente. Sin embargo, después solamente hacemos cero el número (−5) abajo del primer 1 del segundo renglón:

$$\xrightarrow{A_{2,3}(5)} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 9 \\ 0 & 1 & 2 & 4 \\ 0 & 0 & -1 & -3 \end{array}\right) \xrightarrow{M_3(-1)} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 9 \\ 0 & 1 & 2 & 4 \\ 0 & 0 & 1 & 3 \end{array}\right)$$

Ahora, tanto la matriz aumentada del sistema (como la matriz de coeficientes) están en forma escalonada y vemos inmediatamente que $x_3 = 3$. Usaremos entonces *sustitución hacia atrás* para despejar x_2 y luego x_1. La segunda ecuación es $x_2 + 2x_3 = 4$. Así, $x_2 + 2(3) = 4$ y $x_2 = -2$. De igual forma, de la primera ecuación obtenemos $x_1 + 2(-2) + 3(3) = 9$ ó $x_1 = 4$. Por consiguiente hemos obtenido de nuevo la solución (4, −2, 3). Este método de solución se llama *eliminación gaussiana*.

Tenemos, por tanto, dos métodos para resolver nuestros sistemas de ecuaciones:

i. Eliminación de Gauss-Jordan:
Reducir la matriz de coeficientes a la forma escalonada reducida.

ii. Eliminación gaussiana
Reducir la matriz de coeficientes a la forma escalonada, despejar la última incógnita y luego usar sustitución hacia atrás para despejar las otras incógnitas.

¿Cuál de estos dos métodos es más útil? Depende. Si se resuelven sistemas de ecuaciones en una computadora la eliminación gaussiana es el método más usado, ya que requiere menos operaciones elementales sobre renglones. Discutiremos la solución numérica de sistemas de ecuaciones en las Secciones 7.2 y 7.3. Por otro lado, hay ocasiones en las que es necesario obtener la forma escalonada reducida de una matriz (en caso de éstos es discutido en la Sección 1.8). En estos casos el método usado es el de la eliminación de Gauss-Jordan.

Veamos ahora la solución de un sistema general de m ecuaciones en n incógnitas. Dado que lo necesitaremos en la Sección 1.8, resolveremos la mayoría de los sistemas por la eliminación de Gauss-Jordan. Tengamos en mente, sin embargo, que la eliminación gaussiana es, en ocasiones, el método preferido.

En general, un sistema de $m \times n$, de m ecuaciones lineales en n incógnitas está dado por:

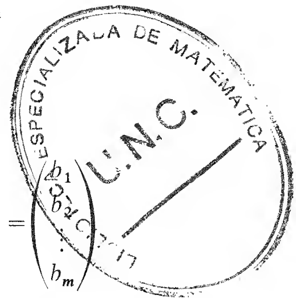
$$\begin{array}{l} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \ldots + a_{2n}x_n = b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \ldots + a_{3n}x_n = b_3 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \ldots + a_{mn}x_n = b_m \end{array} \quad (7)$$

El sistema (7) puede ser escrito en la forma

$$A\mathbf{x} = \mathbf{b}$$

donde

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad \text{y} \quad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$



En este sistema todas las a y b son números reales dados. El problema es encontrar todos los conjuntos de n números, denotados por $(x_1, x_2, x_3, \ldots, x_n)$, que satisfagan cada una de las m ecuaciones de (7). El número a_{ij} es el coeficiente de la variable x_j en la i-ésima ecuación.

Resolvemos el sistema (7) escribiendo el sistema como una matriz aumentada y reduciendo por renglones la matriz a su forma escalonada reducida. Empezamos dividiendo el primer renglón entre a_{11} (operación elemental de renglón (i)). Si $a_{11} = 0$, entonces podemos intercambiar* las ecuaciones de forma que la nueva a_{11} sea diferente de cero. Usamos después esta primera ecuación para eliminar los términos en x_1 de cada una de las otras ecuaciones (usando la operación elemental de renglón (ii)). Luego, la segunda nueva ecuación se divide entre el nuevo término a_{22} y esta nueva ecuación se usa para eliminar los términos en x_2 de las demás ecuaciones. Este proceso se continúa hasta que ocurra una de las tres situaciones siguientes:

i. La última ecuación diferente de cero† queda $x_n = c$ para alguna constante c. Entonces hay una única solución o hay un número infinito de soluciones para el sistema.

ii. La última ecuación diferente de cero queda $a'_{ij}x_j + a'_{i,j+1}x_{j+1} + \cdots + a'_{i,j+k}x_n = c$ para alguna constante c donde al menos dos de las a son diferentes de cero. Esto es, la última ecuación es una ecuación lineal en dos o más de las variables. Entonces existe un número infinito de soluciones.

---

* Para intercambiar un sistema de ecuaciones simplemente escribimos las mismas ecuaciones en diferente orden. Por ejemplo, la primera ecuación puede convertirse en la cuarta ecuación, la tercera ecuación puede convertirse en la segunda ecuación, etc. Esta es una serie de operaciones elementales de renglón (iii).

† La "ecuación cero" es la ecuación 0 = 0.

iii. La última ecuación queda $0 = c$ donde $c \neq 0$. Entonces no existe solución. En este caso decimos que el sistema es *inconsistente*. En los casos (i) y (ii) se dice que el sistema es *consistente*.

Ejemplo 7 Resuelva el sistema

$$x_1+3x_2-5x_3+\ x_4=4$$
$$2x_1+5x_2-2x_3+4x_4=6$$

Solución Escribimos el sistema como una matriz aumentada y la reducimos por renglón:

$$\left(\begin{array}{cccc|c} 1 & 3 & -5 & 1 & 4 \\ 2 & 5 & -2 & 4 & 6 \end{array}\right) \xrightarrow{A_{1,2}(-2)} \left(\begin{array}{cccc|c} 1 & 3 & -5 & 1 & 4 \\ 0 & -1 & 8 & 2 & -2 \end{array}\right) \xrightarrow{M_2(-1)}$$

$$\left(\begin{array}{cccc|c} 1 & 3 & -5 & 1 & 4 \\ 0 & 1 & -8 & -2 & 2 \end{array}\right) \xrightarrow{A_{2,1}(-3)} \left(\begin{array}{cccc|c} 1 & 0 & 19 & 7 & -2 \\ 0 & 1 & -8 & -2 & 2 \end{array}\right)$$

Esto es todo lo que podemos hacer. La matriz de coeficientes está en forma escalonada reducida (caso (*ii*) anterior). Evidentemente hay un número infinito de soluciones. Las variables x_3 y x_4 pueden ser escogidas arbitrariamente. Entonces, $x_2 = 2 + 8x_3 + 2x_4$ y $x_1 = -2 - 19x_3 - 7x_4$. Por lo tanto todas las soluciones están representadas por $(-2 - 19x_3 - 7x_4,\ 2 + 8x_3 + 2x_4,\ x_3,\ x_4)$. Por ejemplo, si $x_3 = 1$ y $x_4 = 2$, obtenemos la solución $(-35, 14, 1, 2)$.

Como se verá después de resolver muchos sistemas, los cálculos se pueden convertir en algo muy complicado. Es una buena regla usar calculadora cuando las fracciones se vuelven muy complejas. Sin embargo, es necesario notar que si los cálculos se llevan a cabo en una calculadora o en una computadora, se pueden introducir errores por "redondeo". Este problema se discutirá en la Sección 7.1.

Concluiremos esta sección con tres ejemplos que ilustran cómo un sistema de ecuaciones lineales se puede presentar en situaciones prácticas.

Ejemplo 8 Un modelo usado frecuentemente en Economía es el *modelo de insumo-producto de Leontief.** Supongamos que un sistema económico tiene n industrias. Cada industria tiene dos tipos de demanda. Primero está la demanda *externa,* que viene de afuera del sistema. Si el sistema es un país, por ejemplo, entonces

* Llamado así en honor del economista estadunidense Wassily W. Leontief. Este modelo fue usado en su publicación "Qualitative Input and Output Relations in the Economic System of the United States" en *Review of Economic Statistics* 18(1936): 105-125. Una versión actualizada de este modelo aparece en el libro de Leontief *Input-Output Analysis* (Nueva York: Oxford University Press, 1966). Leontief obtuvo el premio Nobel de Economía en 1973 por su desarrollo de este análisis.

la demanda externa podría verse como la demanda hecha por otro país. Segundo, hay una demanda hecha a una industria por otra del mismo sistema. En los Estados Unidos, por ejemplo, hay una demanda de productos de la industria acerera hecha por la industria automotriz.

Usemos e_i para representar la demanda externa hecha a la i-ésima industria y a_{ij} para representar la demanda interna hecha a la i-ésima industria por la j-ésima industria. Más precisamente, a_{ij} representa el número de unidades del producto de la industria i necesario para producir una unidad del producto de la industria j. Hagamos que x_i represente la producción de la industria i. Supongamos que la producción de cada industria es igual a su demanda (esto es, no hay sobreproducción). La demanda total es igual a la suma de las demandas interna y externa. Para calcular la demanda interna de la industria 2, por ejemplo, notamos que la industria 1 necesita a_{21} unidades producidas por la industria 2 para generar una unidad de su producción. Si lo que produce la 1 es x_1 entonces $a_{21}x_1$ es la demanda total a la industria 2 por la industria 1. Así, la demanda interna total de la industria 2 es $a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n$.

Llegamos así al siguiente sistema de ecuaciones obtenido de igualar la demanda total con la producción de cada industria:

$$\begin{array}{c} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n+e_1=x_1 \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n+e_2=x_2 \\ \vdots \\ a_{n1}x_1+a_{n2}x_2+\cdots+a_{nn}x_n+e_n=x_n \end{array} \qquad (8)$$

O, escribiendo (8) en la forma del sistema (7), obtenemos

$$\begin{array}{c} (1-a_{11})x_1-\quad a_{12}x_2-\cdots-\quad a_{1n}x_n=e_1 \\ -a_{21}x_1+(1-a_{22})x_2-\cdots-\quad a_{2n}x_n=e_2 \\ \vdots \\ -a_{n1}x_1-\quad a_{n2}x_2-\cdots+(1-a_{nn})x_n=e_n \end{array} \qquad (9)$$

El sistema (9) de n ecuaciones en n incógnitas es muy importante en el análisis económico.

Ejemplo 9 En un sistema económico con tres industrias supongamos que las demandas externas son, respectivamente, 10, 25 y 20. Supongamos que $a_{11}=0.2$, $a_{12}=0.5$, $a_{13}=0.15$, $a_{21}=0.4$, $a_{22}=0.1$, $a_{23}=0.3$, $a_{31}=0.25$, $a_{32}=0.5$ y $a_{33}=0.15$. Encuentre la producción de cada industria de forma que su oferta iguale a su demanda.

Solución Aquí $n = 3$, $1 - a_{11} = 0.8$, $1 - a_{22} = 0.9$ y $1 - a_{33} = 0.85$. Entonces el sistema (9) es

$$0.8x_1-0.5x_2-0.15x_3=10$$
$$-0.4x_1+0.9x_2-\ 0.3x_3=25$$
$$-0.25x_1-0.5x_2+0.85x_3=20$$

Usando calculadora para resolver este sistema obtenemos sucesivamente (con cinco cifras decimales y eliminación de Gauss-Jordan)

$$\left(\begin{array}{rrr|r} 0.8 & -0.5 & -0.15 & 10 \\ -0.4 & 0.9 & -0.3 & 25 \\ -0.25 & -0.5 & 0.85 & 20 \end{array}\right) \xrightarrow{M_1(\frac{1}{0.8})} \left(\begin{array}{rrr|r} 1 & -0.625 & -0.1875 & 12.5 \\ -0.4 & 0.9 & -0.3 & 25 \\ -0.25 & -0.5 & 0.85 & 20 \end{array}\right)$$

$$\xrightarrow[A_{1,3}(0.25)]{A_{1,2}(0.4)} \left(\begin{array}{rrr|r} 1 & -0.625 & -0.1875 & 12.5 \\ 0 & 0.65 & -0.375 & 30 \\ 0 & -0.65625 & 0.80313 & 23.125 \end{array}\right)$$

$$\xrightarrow{M_2(\frac{1}{0.65})} \left(\begin{array}{rrr|r} 1 & -0.625 & -0.1875 & 12.5 \\ 0 & 1 & -0.57692 & 46.15385 \\ 0 & -0.65625 & 0.80313 & 23.125 \end{array}\right)$$

$$\xrightarrow[A_{2,3}(0.65625)]{A_{2,1}(0.625)} \left(\begin{array}{rrr|r} 1 & 0 & -0.54808 & 41.34616 \\ 0 & 1 & -0.57692 & 46.15385 \\ 0 & 0 & 0.42453 & 53.41346 \end{array}\right)$$

$$\xrightarrow{M_3(1/0.42453)} \left(\begin{array}{rrr|r} 1 & 0 & -0.54808 & 41.34616 \\ 0 & 1 & -0.57692 & 46.15385 \\ 0 & 0 & 1 & 125.81787 \end{array}\right)$$

$$\xrightarrow[A_{3,2}(0.57692)]{A_{3,1}(0.54808)} \left(\begin{array}{rrr|r} 1 & 0 & 0 & 110.30442 \\ 0 & 1 & 0 & 118.74070 \\ 0 & 0 & 1 & 125.81787 \end{array}\right)$$

Concluimos que las producciones necesarias para igualar la oferta y la demanda son, aproximadamente, $x_1 = 110$, $x_2 = 119$ y $x_3 = 126$.

Ejemplo 10 Un departamento de Caza y Pesca Estatal suministra tres tipos de alimento a un lago que mantiene a tres especies de peces. Cada pez de la Especie 1 consume cada semana, un promedio de una unidad del Alimento 1, una unidad del Alimento 2 y dos unidades del Alimento 3. Cada pez de la Especie 2, consume cada semana un promedio de tres unidades del Alimento 1, cuatro unidades del Alimento 2 y cinco unidades del Alimento 3. Para un pez de la Especie 3, el consumo semanal promedio es dos unidades del Alimento 1, una unidad del Alimento 2 y cinco unidades del Alimento 3. Cada semana se proporcionan al lago 25,000 unidades del Alimento 1, 20,000 unidades del Alimento 2 y 55,000 unidades del Alimento 3. Si se supone que todo el alimento es ingerido, ¿cuántos peces de cada especie pueden coexistir en el lago?

Solución Si denotamos por x_1, x_2 y x_3 el número de peces de cada una de las tres especies que son mantenidas por el sistema lacustre, utilizando la información en el problema vemos que x_1 peces de la Especie 1 consumen x_1 unidades del Alimento 1, x_2 peces de la Especie 2 consumen $3x_2$ unidades del Alimento 1, y x_3 peces de la Especie 3 consumen $2x_3$ unidades del Alimento 1. Así pues, $x_1 + 3x_2 + 2x_3 = 25{,}000$ = total de suministro semanal del Alimento 1. Obteniendo una ecuación similar para cada uno de los otros dos alimentos, resulta el siguiente sistema:

$$x_1 + 3x_2 + 2x_3 = 25{,}000$$
$$x_1 + 4x_2 + x_3 = 20{,}000$$
$$2x_1 + 5x_2 + 5x_3 = 55{,}000$$

Resolviendo,

$$\left(\begin{array}{rrr|r} 1 & 3 & 2 & 25{,}000 \\ 1 & 4 & 1 & 20{,}000 \\ 2 & 5 & 5 & 55{,}000 \end{array}\right)$$

$$\xrightarrow[A_{1,3}(-2)]{A_{1,2}(-1)} \left(\begin{array}{rrr|r} 1 & 3 & 2 & 25{,}000 \\ 0 & 1 & -1 & -5{,}000 \\ 0 & -1 & 1 & 5{,}000 \end{array}\right) \xrightarrow[A_{2,3}(1)]{A_{2,1}(-3)} \left(\begin{array}{rrr|r} 1 & 0 & 5 & 40{,}000 \\ 0 & 1 & -1 & -5{,}000 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

De este modo, si x_3 se elige en forma arbitraria, tenemos una infinidad de soluciones dadas por $(40{,}000 - 5x_3,\ x_3 - 5{,}000,\ x_3)$. Desde luego, debemos cumplir con las desigualdades $x_1 \geq 0$, $x_2 \geq 0$ y $x_3 \geq 0$. Como $x_2 = x_3 - 5{,}000 \geq 0$, tenemos que $x_3 \geq 5{,}000$. Esto significa que $0 \leq x_1 \leq 40{,}000 - 5(5{,}000) = 15{,}000$. Finalmente, ya que $40{,}000 - 5x_3 \geq 0$, vemos que $x_3 \leq 8{,}000$. Esto significa que las poblaciones que pueden ser mantenidas por el lago son

$$x_1 = 40{,}000 - 5x_3$$
$$x_2 = x_3 - 5{,}000$$
$$5{,}000 \leq x_3 \leq 8{,}000$$

Por ejemplo, si $x_3 = 6{,}000$, entonces $x_1 = 10{,}000$ y $x_2 = 1{,}000$.

Problemas 1.6

En los Problemas del 1 al 20 use la eliminación gaussiana y la eliminación de Gauss-Jordan para encontrar todas las soluciones, si existen, de los sistemas dados.

1. $x_1 - 2x_2 + 3x_3 = 11$
$4x_1 + x_2 - x_3 = 4$
$2x_1 - x_2 + 3x_3 = 10$

2. $-2x_1 + x_2 + 6x_3 = 18$
$5x_1 + 8x_3 = -16$
$3x_1 + 2x_2 - 10x_3 = -3$

3. $3x_1 + 6x_2 - 6x_3 = 9$
$2x_1 - 5x_2 + 4x_3 = 6$
$-x_1 + 16x_2 - 14x_3 = -3$

4. $3x_1 + 6x_2 - 6x_3 = 9$
$2x_1 - 5x_2 + 4x_3 = 6$
$5x_1 + 28x_2 - 26x_3 = -8$

5. $x_1 + x_2 - x_3 = 7$
$4x_1 - x_2 + 5x_3 = 4$
$2x_1 + 2x_2 - 3x_3 = 0$

6. $x_1 + x_2 - x_3 = 7$
$4x_1 - x_2 + 5x_3 = 4$
$6x_1 + x_2 + 3x_3 = 18$

7. $x_1 + x_2 - x_3 = 7$
$4x_1 - x_2 + 5x_3 = 4$
$6x_1 + x_2 + 3x_3 = 20$

8. $x_1 - 2x_2 + 3x_3 = 0$
$4x_1 + x_2 - x_3 = 0$
$2x_1 - x_2 + 3x_3 = 0$

9. $$\begin{aligned} x_1+x_2-x_3&=0\\ 4x_1-x_2+5x_3&=0\\ 6x_1+x_2+3x_3&=0 \end{aligned}$$

10. $$\begin{aligned} 2x_2+5x_3&=6\\ x_1\qquad -2x_3&=4\\ 2x_1+4x_2\qquad&=-2 \end{aligned}$$

11. $$\begin{aligned} x_1+2x_2-x_3&=4\\ 3x_1+4x_2-2x_3&=7 \end{aligned}$$

12. $$\begin{aligned} x_1+2x_2-4x_3&=4\\ -2x_1-4x_2+8x_3&=-8 \end{aligned}$$

13. $$\begin{aligned} x_1+2x_2-4x_3&=4\\ -2x_1-4x_2+8x_3&=-9 \end{aligned}$$

14. $$\begin{aligned} x_1+2x_2-x_3+x_4&=7\\ 3x_1+6x_2-3x_3+3x_4&=21 \end{aligned}$$

15. $$\begin{aligned} 2x_1+6x_2-4x_3+2x_4&=4\\ x_1\qquad -x_3+x_4&=5\\ -3x_1+2x_2-2x_3\qquad&=-2 \end{aligned}$$

16. $$\begin{aligned} x_1-2x_2+x_3+x_4&=2\\ 3x_1\qquad +2x_3-2x_4&=-8\\ 4x_2-x_3-x_4&=1\\ -x_1+6x_2-2x_3\qquad&=7 \end{aligned}$$

17. $$\begin{aligned} x_1-2x_2+x_3+x_4&=2\\ 3x_1\qquad +2x_3-2x_4&=-8\\ 4x_2-x_3-x_4&=1\\ 5x_1\qquad +3x_3-x_4&=-3 \end{aligned}$$

18. $$\begin{aligned} x_1-2x_2+x_3+x_4&=2\\ 3x_1\qquad +2x_3-2x_4&=-8\\ 4x_2-x_3-x_4&=1\\ 5x_1\qquad +3x_3-x_4&=0 \end{aligned}$$

19. $$\begin{aligned} x_1+x_2&=4\\ 2x_1-3x_2&=7\\ 3x_1+2x_2&=8 \end{aligned}$$

20. $$\begin{aligned} x_1+x_2&=4\\ 2x_1-3x_2&=7\\ 3x_1-2x_2&=11 \end{aligned}$$

En los Problemas del 21 *al* 29 *determine si la matriz dada está en forma escalonada (pero no en forma escalonada reducida), en forma escalonada reducida o en ninguna de las dos.*

21. $\begin{pmatrix}1&1&0\\0&1&1\\0&0&1\end{pmatrix}$ **22.** $\begin{pmatrix}2&0&0\\0&1&0\\0&0&-1\end{pmatrix}$ **23.** $\begin{pmatrix}1&0&1&0\\0&1&1&0\\0&0&0&0\end{pmatrix}$

24. $\begin{pmatrix}1&0&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix}$ **25.** $\begin{pmatrix}0&1&0&0\\1&0&0&0\\0&0&0&0\end{pmatrix}$ **26.** $\begin{pmatrix}1&0&1&2\\0&1&3&4\end{pmatrix}$

27. $\begin{pmatrix}1&0\\0&1\\0&0\end{pmatrix}$ **28.** $\begin{pmatrix}1&0&0\\0&0&0\\0&0&1\end{pmatrix}$ **29.** $\begin{pmatrix}1&0&0&4\\0&1&0&5\\0&1&1&6\end{pmatrix}$

En los Problemas del 30 *al* 35 *use operaciones elementales de renglón para reducir las matrices dadas a forma escalonada y a forma escalonada reducida.*

30. $\begin{pmatrix}1&1\\2&3\end{pmatrix}$ **31.** $\begin{pmatrix}-1&6\\4&2\end{pmatrix}$ **32.** $\begin{pmatrix}1&-1&1\\2&4&3\\5&6&-2\end{pmatrix}$

33. $\begin{pmatrix}2&-4&8\\3&5&8\\-6&0&4\end{pmatrix}$ **34.** $\begin{pmatrix}2&-4&-2\\3&1&6\end{pmatrix}$ **35.** $\begin{pmatrix}2&-7\\3&5\\4&-3\end{pmatrix}$

36. En el modelo de Leontief de insumo-producto del Ejemplo 8 suponga que hay tres industrias. Suponga además que $e_1 = 10$, $e_2 = 15$, $e_3 = 30$, $a_{11} = \frac{1}{3}$, $a_{12} = \frac{1}{2}$, $a_{13} = \frac{1}{6}$, $a_{21} = \frac{1}{4}$, $a_{22} = \frac{1}{4}$, $a_{23} = \frac{1}{8}$, $a_{31} = \frac{1}{12}$, $a_{32} = \frac{1}{3}$ y $a_{33} = \frac{1}{6}$. Encuentre la producción de cada industria suponiendo que la oferta es igual a la demanda.

37. En el Ejemplo 10, suponga que se suministran al lago 15,000 unidades del primer alimento, 10,000 del segundo y 35,000 del tercero. Suponiendo que los tres alimentos se consumen, ¿qué población de cada especie se encontrará en coexistencia? ¿Existe una única solución?

38. Un viajero recién regresado de Europa gastó en alojamiento, por día, \$30 dólares en Inglaterra, \$20 en Francia y \$20 en España. En comidas, por día, gastó \$20 en Inglaterra, \$30 en Francia y \$20 en España. Adicionalmente, desembolsó \$10 por día en cada país en gastos varios. El registro de nuestro viajero indica que gastó un total de \$340 en alojamiento, \$320 en alimentación y \$140 en gastos varios en su recorrido por estos tres países. Calcule el número de días que permaneció el viajero en cada país o muestre que el registro debe ser incorrecto, pues las cantidades gastadas son incompatibles entre sí.

39. Una inversionista le afirma a un corredor de bolsa, que todas sus acciones son de tres empresas, una de aviación, una de hoteles y una de alimentos, y que hace 2 días que el precio de sus acciones bajó en \$350 pero que subió a \$600 el día de ayer. El corredor recuerda que hace dos días el precio por acción de la compañía de aviación bajó \$1, que el de la compañía de hoteles bajó \$1.50, pero que el de la de los alimentos subió \$0.50. También recordó el corredor que el día de ayer los precios por acción de las tres compañías se comportaron como sigue. Subió el de aviación \$1.50, el de hoteles cayó \$0.50, y el de la de alimentos subió \$1. Muestre que el corredor no posee información suficiente para calcular el número de acciones de las que es dueña la inversionista, pero que cuando ella diga que tiene 200 acciones de la empresa alimenticia, entonces sí es posible determinar el número de acciones restante.

40. Un espía sabe que están estacionados 60 aeroplanos —aviones de caza y bombarderos— en un cierto aeropuerto secreto. El agente quiere determinar cuántos aeroplanos hay de cada clase. Existe un cierto tipo de cohete que es portado por ambas clases de aviones; el avión caza porta seis de estos cohetes y el bombardero dos. El agente averigua que se requieren 250 cohetes para poder armar a todos los aeroplanos del aeródromo secreto. Más aún, el agente oye que existen dos veces más aviones cazas que bombarderos en la base aérea en cuestión (esto es, el número de aviones cazas menos dos veces el número de bombarderos da cero). Calcule el número de interceptores y de bombarderos o muestre que la información del agente es incorrecta, por ser incompatible.

41. Considere el sistema

$$\begin{aligned} 2x_1-x_2+3x_3&=a\\ 3x_1+x_2-5x_3&=b\\ -5x_1-5x_2+21x_3&=c \end{aligned}$$

Muestre que el sistema es inconsistente si $c \neq 2a - 3b$.

42. Considere el sistema

$$\begin{aligned} 2x_1+3x_2-x_3&=a\\ x_1-x_2+3x_3&=b\\ 3x_1+7x_2-5x_3&=c \end{aligned}$$

Encuentre condiciones en a, b y c de forma que el sistema sea consistente.

★ **43.** Considere el sistema general de tres ecuaciones lineales en tres incógnitas:

$$\begin{aligned} a_{11}x_1+a_{12}x_2+a_{13}x_3&=b_1\\ a_{21}x_1+a_{22}x_2+a_{23}x_3&=b_2\\ a_{31}x_1+a_{32}x_2+a_{33}x_3&=b_3 \end{aligned}$$

Encuentre condiciones en los coeficientes a_{ij} de forma que el sistema tenga una única solución.

44. Resuelva el siguiente sistema usando una calculadora e incluyendo cinco cifras decimales de exactitud:

$$\begin{aligned} 2x_2 - x_3 - 4x_4 &= 2 \\ x_1 - x_2 + 5x_3 + 2x_4 &= -4 \\ 3x_1 + 3x_2 - 7x_3 - x_4 &= 4 \\ -x_1 - 2x_2 + 3x_3 \qquad &= -7 \end{aligned}$$

45. Haga lo mismo para el sistema

$$\begin{aligned} 3.8x_1 + 1.6x_2 + 0.9x_3 &= 3.72 \\ -0.7x_1 + 5.4x_2 + 1.6x_3 &= 3.16 \\ 1.5x_1 + 1.1x_2 - 3.2x_3 &= 43.78 \end{aligned}$$

En los Problemas del 46 al 51 escriba el sistema dado en la forma $A\mathbf{x} = \mathbf{b}$.

46. $\begin{aligned} 2x_1 - x_2 &= 3 \\ 4x_1 + 5x_2 &= 7 \end{aligned}$

47. $\begin{aligned} x_1 - x_2 + 3x_3 &= 11 \\ 4x_1 + x_2 - x_3 &= -4 \\ 2x_1 - x_2 + 3x_3 &= 10 \end{aligned}$

48. $\begin{aligned} 3x_1 + 6x_2 - 7x_3 &= 0 \\ 2x_1 - x_2 + 3x_3 &= 1 \end{aligned}$

49. $\begin{aligned} 4x_1 - x_2 + x_3 - x_4 &= -7 \\ 3x_1 + x_2 - 5x_3 + 6x_4 &= 8 \\ 2x_1 - x_2 + x_3 \qquad &= 9 \end{aligned}$

50. $\begin{aligned} x_2 - x_3 &= 7 \\ x_1 \qquad + x_3 &= 2 \\ 3x_1 + 2x_2 \qquad &= -5 \end{aligned}$

51. $\begin{aligned} 2x_1 + 3x_2 - x_3 &= 0 \\ -4x_1 + 2x_2 + x_3 &= 0 \\ 7x_1 + 3x_2 - 9x_3 &= 0 \end{aligned}$

1.7 Sistemas homogéneos de ecuaciones

El sistema general de ecuaciones lineales de $m \times n$, sistema (1.6.7), se conoce como *homogéneo* si todas las constantes $b_1, b_2, \ldots, b_m$ son cero. Esto es, el sistema homogéneo general está dado por

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0 \\ \vdots \qquad \vdots \qquad\quad \vdots \quad & \;\; \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= 0 \end{aligned} \tag{1}$$

Los sistemas homogéneos surgen de varias maneras. Veremos una de éstas en la Sección 4.4. En esta sección resolveremos algunos sistemas homogéneos, usando nuevamente el método de eliminación de Gauss-Jordan.

Para el sistema general lineal hay tres posibilidades: no hay solución, hay una o hay un número infinito de soluciones. Para el sistema general homogéneo la situación es más simple. Como $x_1 = x_2 = \cdots = x_n = 0$ siempre es una solución (llamada la *solución trivial* o la *solución cero*), solamente hay dos posibilidades: la solución cero es la única solución o hay un número infinito de soluciones además de la solución cero. (Las soluciones distintas de la solución cero se conocen como las *soluciones no triviales*).

Observamos que (1) se puede expresar como

$$A\mathbf{x} = \mathbf{0}$$

en donde

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad \text{y} \quad \mathbf{0} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \leftarrow m \text{ ceros}$$

Ejemplo 1 Resuelva el sistema homogéneo

$$\begin{aligned} 2x_1 + 4x_2 + 6x_3 &= 0 \\ 4x_1 + 5x_2 + 6x_3 &= 0 \\ 3x_1 + x_2 - 2x_3 &= 0 \end{aligned}$$

Solución Esta es la versión homogénea del sistema del Ejemplo 1.6.1. Puede ser escrito en la forma $A\mathbf{x} = \mathbf{0}$ donde A y $\mathbf{x}$ son como en el ejemplo citado y

$$\mathbf{0} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Reduciendo en forma sucesiva, se obtiene (después de dividir la primera ecuación entre 2)

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 0 \\ 4 & 5 & 6 & 0 \\ 3 & 1 & -2 & 0 \end{array}\right) \xrightarrow[A_{1,3}(-3)]{A_{1,2}(-4)} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 0 \\ 0 & -3 & -6 & 0 \\ 0 & -5 & -11 & 0 \end{array}\right) \xrightarrow{M_2(-\frac{1}{3})} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & -5 & -11 & 0 \end{array}\right)$$

$$\xrightarrow[A_{2,3}(5)]{A_{2,1}(-2)} \left(\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & -1 & 0 \end{array}\right) \xrightarrow{M_3(-1)} \left(\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right) \xrightarrow[A_{3,2}(-2)]{A_{3,1}(1)} \left(\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right)$$

Así, el sistema tiene como única solución al (0,0,0). Esto es, el sistema solamente tiene la solución trivial.

Ejemplo 2 Resuelva el sistema homogéneo

$$\begin{aligned} x_1 + 2x_2 - x_3 &= 0 \\ 3x_1 - 3x_2 + 2x_3 &= 0 \\ -x_1 - 11x_2 + 6x_3 &= 0 \end{aligned}$$

Solución Reduciendo con eliminación de Gauss-Jordan obtenemos, sucesivamente,

$$\left(\begin{array}{ccc|c} 1 & 2 & -1 & 0 \\ 3 & -3 & 2 & 0 \\ -1 & -11 & 6 & 0 \end{array}\right) \xrightarrow[A_{1,3}(1)]{A_{1,2}(-3)} \left(\begin{array}{ccc|c} 1 & 2 & -1 & 0 \\ 0 & -9 & 5 & 0 \\ 0 & -9 & 5 & 0 \end{array}\right)$$

$$\xrightarrow{M_2(-\frac{1}{9})}\left(\begin{array}{ccc|c}1 & 2 & -1 & 0\\ 0 & 1 & -\frac{5}{9} & 0\\ 0 & -9 & 5 & 0\end{array}\right)\xrightarrow[A_{2,3}(9)]{A_{2,1}(-2)}\left(\begin{array}{ccc|c}1 & 0 & \frac{1}{9} & 0\\ 0 & 1 & -\frac{5}{9} & 0\\ 0 & 0 & 0 & 0\end{array}\right)$$

La matriz aumentada está en forma escalonada reducida y, evidentemente, hay un número infinito de soluciones dadas por $(-\frac{1}{9}x_3, \frac{5}{9}x_3, x_3)$. Si $x_3 = 0$, por ejemplo, obtenemos la solución trivial. Si $x_3 = 1$ obtenemos la solución $(-\frac{1}{9}, \frac{5}{9}, 1)$.

Ejemplo 3 Resuelva el sistema

$$\begin{aligned} x_1 + x_2 - x_3 &= 0\\ 4x_1 - 2x_2 + 7x_3 &= 0 \end{aligned} \tag{2}$$

Solución Reduciendo por renglones obtenemos

$$\left(\begin{array}{ccc|c}1 & 1 & -1 & 0\\ 4 & -2 & 7 & 0\end{array}\right)\xrightarrow{A_{1,2}(-4)}\left(\begin{array}{ccc|c}1 & 1 & -1 & 0\\ 0 & -6 & 11 & 0\end{array}\right)$$

$$\xrightarrow{M_2(-\frac{1}{6})}\left(\begin{array}{ccc|c}1 & 1 & -1 & 0\\ 0 & 1 & -\frac{11}{6} & 0\end{array}\right)\xrightarrow{A_{2,1}(-1)}\left(\begin{array}{ccc|c}1 & 0 & \frac{5}{6} & 0\\ 0 & 1 & -\frac{11}{6} & 0\end{array}\right)$$

Así, hay un número infinito de soluciones dadas por $(-\frac{5}{6}x_3, \frac{11}{6}x_3, x_3)$. Esto era de esperarse ya que el sistema (2) tiene tres incógnitas y sólo dos ecuaciones.

De hecho, si hay más incógnitas que ecuaciones, el sistema homogéneo (1) siempre tendrá un número infinito de soluciones. Para ver esto, notemos que si la única solución es la trivial entonces la reducción por renglón nos lleva al sistema

$$\begin{aligned} x_1 \qquad\qquad &= 0\\ x_2 \qquad &= 0\\ &\vdots\\ x_n &= 0 \end{aligned}$$

y, posiblemente, a ecuaciones adicionales de la forma $0 = 0$. Pero este sistema tiene al menos tantas ecuaciones como incógnitas. Debido a que la reducción por renglón no cambia ni el número de ecuaciones ni el número de incógnitas, tenemos una contradicción a nuestra suposición de que hay más incógnitas que ecuaciones. De aquí tenemos el Teorema 1.

Teorema 1 El sistema homogéneo (1) tiene un número infinito de soluciones si $n > m$.

Existe una relación fundamental entre los sistemas homogéneos y los no homogéneos. Sea A una matriz $m \times n$,

$$\mathbf{x} = \begin{pmatrix} x_1\\ x_2\\ \vdots\\ x_n \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} b_1\\ b_2\\ \vdots\\ b_m \end{pmatrix}, \quad \text{y} \quad \mathbf{0} = \begin{pmatrix} 0\\ 0\\ \vdots\\ 0 \end{pmatrix} \ (m \text{ ceros})$$

El sistema general no homogéneo se puede escribir como

$$A\mathbf{x} = \mathbf{b} \tag{3}$$

Con A y $\mathbf{x}$ como en (3), se define el *sistema homogéneo asociado* por

$$A\mathbf{x} = \mathbf{0}. \tag{4}$$

Teorema 2 Sean $\mathbf{x}_1$ y $\mathbf{x}_2$ soluciones del sistema no homogéneo (3). Entonces su diferencia $\mathbf{x}_1 - \mathbf{x}_2$, es una solución del sistema homogéneo asociado (4).

Demostración

por la ley distributiva (7) (en la página 25)

$$A(\mathbf{x}_1 - \mathbf{x}_2) = A\mathbf{x}_1 - A\mathbf{x}_2 = \mathbf{b} - \mathbf{b} = \mathbf{0}. \ \blacksquare$$

Corolario Sea $\mathbf{x}$ una solución particular del sistema no homogéneo (3) y sea $\mathbf{y}$ otra solución de (3). Entonces existe un vector $\mathbf{h}$ que es solución del sistema homogéneo (4) de modo que

$$\mathbf{y} = \mathbf{x} + \mathbf{h} \tag{5}$$

Demostración Si $\mathbf{h}$ se define por $\mathbf{h} = \mathbf{y} - \mathbf{x}$, entonces $\mathbf{h}$ es solución de (4) por el Teorema 2 y la relación (5) es inmediata.

El Teorema 2 y su corolario son muy útiles. Expresan que

> Para encontrar todas las soluciones del sistema no homogéneo (3), es necesario encontrar sólo *una* solución a (3) y todas las soluciones asociadas al sistema (4).

Observación. Un resultado similar vale para las soluciones de ecuaciones diferenciales homogéneas y no homogéneas (véanse los Problemas 23 y 24). Una de las muchas características atractivas de las matemáticas es la cercana interrelación que muestra entre temas aparentemente lejanos.

Ejemplo 4 Usando el resultado anterior, encuentre todas las soluciones del sistema no homogéneo

$$\begin{aligned} x_1 + 2x_2 - x_3 &= 2\\ 2x_1 + 3x_2 + 5x_3 &= 5\\ -x_1 - 3x_2 + 8x_3 &= -1 \end{aligned} \tag{6}$$

Solución Primero se halla una solución usando reducción por renglones:

$$\left(\begin{array}{ccc|c} 1 & 2 & -1 & 2 \\ 2 & 3 & 5 & 5 \\ -1 & -3 & 8 & -1 \end{array}\right) \xrightarrow[A_{1,3}(1)]{A_{1,2}(-2)} \left(\begin{array}{ccc|c} 1 & 2 & -1 & 2 \\ 0 & -1 & 7 & 1 \\ 0 & -1 & 7 & 1 \end{array}\right) \xrightarrow[A_{2,3}(-1)]{A_{2,1}(2)}$$

$$\left(\begin{array}{ccc|c} 1 & 0 & 13 & 4 \\ 0 & -1 & 7 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Vemos que existe un número infinito de soluciones. Haciendo $x_3 = 0$ (se puede usar cualquier otro número) obtenemos $x_1 = 4$ y $x_2 = -1$. De este modo, una solución particular queda dada por $\mathbf{x_p} = (4, -1, 0)$.

La reducción por renglones del sistema homogéneo asociado, nos lleva a

$$\left(\begin{array}{ccc|c} 1 & 0 & 13 & 0 \\ 0 & -1 & 7 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Por lo tanto, todas las soluciones del sistema homogéneo satisfacen las relaciones

$$x_1 = -13x_3, \qquad x_2 = 7x_3$$

o lo que es lo mismo,

$$\mathbf{x_h} = (x_1, x_2, x_3) = (-13x_3, 7x_3, x_3) = x_3(-13, 7, 1)$$

De esta manera podemos escribir cada solución del sistema (6) como

$$\mathbf{x} = \mathbf{x_p} + \mathbf{x_h} = (4, -1, 0) + x_3(-13, 7, 1)$$

sujeta a la elección de un valor apropiado de x_3. Por ejemplo, para $x_3 = 0$, se obtiene la solución $(4, -1, 0)$, mientras que si $x_3 = 2$ resulta $(-22, 13, 2)$.

Problemas 1.7

En los Problemas del 1 *al* 13 *encuentre todas las soluciones a los sistemas homogéneos.*

1. $\begin{aligned} 2x_1 - x_2 &= 0 \\ 3x_1 + 4x_2 &= 0 \end{aligned}$

2. $\begin{aligned} x_1 - 5x_2 &= 0 \\ -x_1 + 5x_2 &= 0 \end{aligned}$

3. $\begin{aligned} x_1 + x_2 - x_3 &= 0 \\ 2x_1 - 4x_2 + 3x_3 &= 0 \\ 3x_1 + 7x_2 - x_3 &= 0 \end{aligned}$

4. $\begin{aligned} x_1 + x_2 - x_3 &= 0 \\ 2x_1 - 4x_2 + 3x_3 &= 0 \\ -x_1 - 7x_2 + 6x_3 &= 0 \end{aligned}$

5. $\begin{aligned} x_1 + x_2 - x_3 &= 0 \\ 2x_1 - 4x_2 + 3x_3 &= 0 \\ -5x_1 + 13x_2 - 10x_3 &= 0 \end{aligned}$

6. $\begin{aligned} 2x_1 + 3x_2 - x_3 &= 0 \\ 6x_1 - 5x_2 + 7x_3 &= 0 \end{aligned}$

7. $\begin{aligned} 4x_1 - x_2 &= 0 \\ 7x_1 + 3x_2 &= 0 \\ -8x_1 + 6x_2 &= 0 \end{aligned}$

8. $\begin{aligned} x_1 - x_2 + 7x_3 - x_4 &= 0 \\ 2x_1 + 3x_2 - 8x_3 + x_4 &= 0 \end{aligned}$

9. $\begin{aligned} x_1 - 2x_2 + x_3 + x_4 &= 0 \\ 3x_1 + 2x_3 - 2x_4 &= 0 \\ 4x_2 - x_3 - x_4 &= 0 \\ 5x_1 + 3x_3 - x_4 &= 0 \end{aligned}$

10. $\begin{aligned} -2x_1 + 7x_4 &= 0 \\ x_1 + 2x_2 - x_3 + 4x_4 &= 0 \\ 3x_1 - x_3 + 5x_4 &= 0 \\ 4x_1 + 2x_2 + 3x_3 &= 0 \end{aligned}$

11. $\begin{aligned} 2x_1 - x_2 &= 0 \\ 3x_1 + 5x_2 &= 0 \\ 7x_1 - 3x_2 &= 0 \\ -2x_1 + 3x_2 &= 0 \end{aligned}$

12. $\begin{aligned} x_1 - 3x_2 &= 0 \\ -2x_1 + 6x_2 &= 0 \\ 4x_1 - 12x_2 &= 0 \end{aligned}$

13. $\begin{aligned} x_1 + x_2 - x_3 &= 0 \\ 4x_1 - x_2 + 5x_3 &= 0 \\ -2x_1 + x_2 - 2x_3 &= 0 \\ 3x_1 + 2x_2 - 6x_3 &= 0 \end{aligned}$

14. Muestre que el sistema homogéneo

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= 0 \\ a_{21}x_1 + a_{22}x_2 &= 0 \end{aligned}$$

tiene un número infinito de soluciones si y sólo si el determinante del sistema es cero.

15. Considere el sistema

$$\begin{aligned} 2x_1 - 3x_2 + 5x_3 &= 0 \\ -x_1 + 7x_2 - x_3 &= 0 \\ 4x_1 - 11x_2 + kx_3 &= 0 \end{aligned}$$

¿Para qué valor de k este sistema tiene soluciones no triviales?

★ **16.** Considere el sistema homogéneo de 3×3

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= 0 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= 0 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= 0 \end{aligned}$$

Encuentre condiciones en los coeficientes a_{ij} de forma que la solución cero sea la única solución.

En los Problemas del 17 *al* 22 *obtenga todas las soluciones a los sistemas no homogéneos dados. Encontrando primero una solución (si es que existe) y después hallando todas las soluciones del sistema homogéneo asociado.*

17. $\begin{aligned} x_1 - 3x_2 &= 2 \\ -2x_1 + 6x_2 &= -4 \end{aligned}$

18. $\begin{aligned} x_1 - x_2 + x_3 &= 6 \\ 3x_1 - 3x_2 + 3x_3 &= 18 \end{aligned}$

19. $\begin{aligned} x_1 - x_2 - x_3 &= 2 \\ 2x_1 + x_2 + 2x_3 &= 4 \\ x_1 - 4x_2 - 5x_3 &= 2 \end{aligned}$

20. $\begin{aligned} x_1 - x_2 - x_3 &= 2 \\ 2x_1 + x_2 + 2x_3 &= 4 \\ x_1 - 4x_2 - 5x_3 &= 2 \end{aligned}$

21. $\begin{aligned} x_1 + x_2 - x_3 + 2x_4 &= 3 \\ 3x_1 + 2x_2 + x_3 - x_4 &= 5 \end{aligned}$

22. $\begin{aligned} x_1 - x_2 + x_3 - x_4 &= -2 \\ -2x_1 + 3x_2 - x_3 + 2x_4 &= 5 \\ 4x_1 - 2x_2 + 2x_3 - 3x_4 &= 6 \end{aligned}$

*□**23.** Considere la ecuación diferencial lineal y homogénea de segundo grado

$$y''(x) + a(x)y'(x) + b(x)y(x) = 0 \tag{7}$$

donde $a(x)$ y $b(x)$ son continuas y de la función incógnita y se supone que posee segunda derivada. Demuestre que si y_1, y_2 son soluciones de (7), entonces $c_1y_1 + c_2y_2$ es también solución para cualesquiera constantes y_1 y y_2.

□**24.** Suponga que y_p y y_q son soluciones del sistema no homogéneo

$$y''(x) + a(x)y'(x) + b(x)y(x) = f(x) \tag{8}$$

Demuestre que $y_p - y_q$ es solución de (7).

* El símbolo □ indica que se requiere cierto conocimiento del Cálculo para resolver el problema.

1.8 La inversa de una matriz cuadrada

En esta sección definimos dos clases de matrices que tienen un papel importante en la teoría de matrices. Empezamos con un ejemplo simple. Sean $A=\begin{pmatrix}2 & 5\\ 1 & 3\end{pmatrix}$ y $B=\begin{pmatrix}3 & -5\\ -1 & 2\end{pmatrix}$. Entonces un cálculo simple muestra que $AB = BA = I_2$, donde $I_2 = \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$. La matriz I_2 se conoce como la *matriz identidad* de 2×2. La matriz B se conoce como la *inversa* de A y se escribe como A^{-1}.

Definición 1 **Matriz identidad.** La *matriz identidad* de $n \times n$ es la matriz de $n \times n$ en la que las componentes de la *diagonal principal** son 1, y 0 en todas las demás posiciones. Esto es,

$$I_n = (b_{ij}) \quad \text{donde} \quad b_{ij} = \begin{cases}1 & \text{si } i=j\\ 0 & \text{si } i\neq j\end{cases} \qquad (1)$$

Ejemplo 1 $I_3=\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}$ y $I_5=\begin{pmatrix}1 & 0 & 0 & 0 & 0\\ 0 & 1 & 0 & 0 & 0\\ 0 & 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 0 & 1\end{pmatrix}$

Teorema 1 Sea A una matriz cuadrada de $n\times n$. Entonces

$$AI_n = I_nA = A$$

Esto es, I_n conmuta con cualquier matriz de $n\times n$ y la deja intacta después de la multiplicación por la izquierda o por la derecha.

Nota. I_n juega el mismo papel para las matrices de $n \times n$, que el número 1 para los números reales (pues $1 \cdot a = a \cdot 1 = a$ para cualquier número real a).

* La diagonal principal de $A = (a_{ij})$ consiste de las componentes a_{11}, a_{22}, a_{33}, etc. A menos que se especifique otra cosa, nos referiremos a la diagonal principal simplemente como la *diagonal*.

Demostración Sea c_{ij} el ij-ésimo elemento de AI_n. Entonces

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ij}b_{jj} + \cdots + a_{in}b_{nj}.$$

Pero, por (1), esta suma es igual a a_{ij}. Así, $AI_n = A$. De forma similar podemos mostrar que $I_nA = A$ y esto prueba el teorema. ■

Notación De aquí en adelante escribiremos la matriz identidad simplemente como I, pues si A es de $n\times n$, los productos IA y AI están definidos sólo si I es también de $n\times n$.

Definición 2 ***Inversa de una matriz.*** Sean A y B matrices de $n \times n$. Supongamos que

$$AB = BA = I$$

Entonces B se conoce como la *inversa* de A y se escribe A^{-1}. Así tenemos

$$AA^{-1} = A^{-1}A = I$$

Si A tiene una inversa entonces decimos que A es *invertible*.

Observación 1. Se sigue inmediatamente de esta definición que $(A^{-1})^{-1} = A$ si A es *invertible*.

Observación 2. Esta definición *no* establece que toda matriz cuadrada tenga una inversa. De hecho hay muchas matrices cuadradas que no tienen inversa. (Vea el Ejemplo 3 siguiente.)

En la Definición 2 definimos *la* inversa de una matriz. Esta afirmación sugiere que las inversas son únicas. En el teorema siguiente veremos que éste es el caso.

Teorema 2 Si una matriz cuadrada A es invertible, entonces su inversa es única.

Demostración Supongamos que B y C son dos inversas para A. Podemos mostrar que $B = C$. Por definición tenemos $AB = BA = I$ y $AC = CA = I$. Entonces $B(AC) = BI = B$ y $(BA)C = IC = C$. Pero $B(AC) = (BA)C$ por la ley asociativa de la multiplicación de matrices. Así, $B = C$ y el teorema está demostrado. ■

Aquí damos otro hecho importante referente a matrices inversas.

Teorema 3 Sean A y B matrices invertibles de $n\times n$. Entonces AB es invertible y

$$(AB)^{-1} = B^{-1}A^{-1}$$

Demostración Para demostrar este resultado nos referimos a la Definición 2. Esto es, $B^{-1}A^{-1} = (AB)^{-1}$ si y sólo si $B^{-1}A^{-1}(AB) = (AB)(B^{-1}A^{-1}) = I$. Pero esto se desprende puesto que

Ecuación (6) en la página 25

$$(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}IB = B^{-1}B = I$$

y

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I. \blacksquare$$

Considere el sistema de n ecuaciones en n incógnitas

$$A\mathbf{x} = \mathbf{b},$$

y suponga que A es invertible. Entonces

$$A^{-1}A\mathbf{x} = A^{-1}\mathbf{b} \quad \text{(si multiplicamos el primer miembro por } A^{-1})$$

$$I\mathbf{x} = A^{-1}\mathbf{b} \qquad A^{-1}A = I$$

$$\mathbf{x} = A^{-1}\mathbf{b} \qquad I\mathbf{x} = \mathbf{x}$$

Esto es,

> Si A es invertible, el sistema $A\mathbf{x} = \mathbf{b}$ tiene la solución única $\mathbf{x} = A^{-1}\mathbf{b}$. (2)

Esta es una de las razones por las que se estudian las inversas de matrices.

Una vez definida la inversa de una matriz, ocurren dos preguntas en forma inmediata:

1. ¿Qué matrices son aquéllas que poseen inversa?
2. Si una matriz tiene inversa, ¿de qué modo la podemos evaluar?

Contestaremos ambas preguntas en esta sección. Antes de dar lo que pudiera parecer un conjunto de reglas arbitrarias, examinaremos el caso de 2 × 2.

Ejemplo 2 Sea $A = \begin{pmatrix} 2 & -3 \\ -4 & 5 \end{pmatrix}$. Calcular A^{-1}, si existe.

Solución Supóngase que A^{-1} existe. Escribimos $A^{-1} = \begin{pmatrix} x & y \\ z & w \end{pmatrix}$ y utilizamos el hecho de que $AA^{-1} = I$. Entonces

$$AA^{-1} = \begin{pmatrix} 2 & -3 \\ -4 & 5 \end{pmatrix}\begin{pmatrix} x & y \\ z & w \end{pmatrix} = \begin{pmatrix} 2x - 3z & 2y - 3w \\ -4x + 5z & -4y + 5w \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Las últimas dos matrices sólo pueden ser iguales si sus componentes correspondientes son asimismo iguales. Esto significa que

$$2x - 3z = 1 \qquad (3)$$

$$2y - 3w = 0 \qquad (4)$$

$$-4x + 5z = 0 \qquad (5)$$

$$-4y + 5w = 1 \qquad (6)$$

Este es un sistema de cuatro ecuaciones con cuatro incógnitas. Nótese que existen dos ecuaciones que contienen a x y z solamente [ecuaciones (3) y (5)] y dos ecuaciones que contienen las variables y, w [ecuaciones (4) y (6)]. Escribimos estos dos sistemas en la forma de matriz aumentada:

$$\left(\begin{array}{cc|c} 2 & -3 & 1 \\ -4 & 5 & 0 \end{array}\right) \qquad (7)$$

$$\left(\begin{array}{cc|c} 2 & -3 & 0 \\ -4 & 5 & 1 \end{array}\right) \qquad (8)$$

Ahora bien, sabemos de la Sección 1.6 que si el sistema (7) (en las variables x, z) tiene una solución única, entonces la eliminación de Gauss-Jordan de (7) resulta en

$$\left(\begin{array}{cc|c} 1 & 0 & x \\ 0 & 1 & z \end{array}\right)$$

en donde (x, z) es el par único de números que satisface $2x - 3z = 1$ y $-4x + 5z = 0$. Similarmente, la reducción por renglones de (8) conduce a

$$\left(\begin{array}{cc|c} 1 & 0 & y \\ 0 & 1 & w \end{array}\right)$$

donde (y, w) es el par único de números que satisface $2y - 3w = 0$ y simultáneamente, $-4y + 5w = 1$.

Como las matrices de coeficientes en (7) y (8) son las mismas, podemos realizar simultáneamente la reducción por renglones en las dos matrices aumentadas, utilizando la nueva matriz aumentada

$$\left(\begin{array}{cc|cc} 2 & -3 & 1 & 0 \\ -4 & 5 & 0 & 1 \end{array}\right) \qquad (9)$$

Si A^{-1} es invertible, entonces el sistema definido por (3), (4), (5) y (6) tiene una solución única, y por lo dicho anteriormente, al efectuar la eliminación de Gauss-Jordan resulta que

$$\left(\begin{array}{cc|cc} 1 & 0 & x & y \\ 0 & 1 & z & w \end{array}\right)$$

Ahora se realizan los cálculos notando que la matriz de la izquierda en (9) es A y que la matriz de la derecha en (9) es I:

$$\left(\begin{array}{cc|cc} 2 & -3 & 1 & 0 \\ -4 & 5 & 0 & 1 \end{array}\right) \xrightarrow{M_1(\frac{1}{2})} \left(\begin{array}{cc|cc} 1 & -\frac{3}{2} & \frac{1}{2} & 0 \\ -4 & 5 & 0 & 1 \end{array}\right)$$

$$\xrightarrow{A_{1,2}(4)} \left(\begin{array}{cc|cc} 1 & -\frac{3}{2} & \frac{1}{2} & 0 \\ 0 & -1 & 2 & 1 \end{array}\right)$$

$$\xrightarrow{M_2(-1)} \left(\begin{array}{cc|cc} 1 & -\frac{3}{2} & \frac{1}{2} & 0 \\ 0 & 1 & -2 & -1 \end{array}\right)$$

$$\xrightarrow{A_{2,1}(\frac{3}{2})} \left(\begin{array}{cc|cc} 1 & 0 & -\frac{5}{2} & -\frac{3}{2} \\ 0 & 1 & -2 & -1 \end{array}\right)$$

Así pues, $x = -\frac{5}{2}$, $y = -\frac{3}{2}$, $z = -2$, $w = -1$, y $A^{-1} = \begin{pmatrix} -\frac{5}{2} & -\frac{3}{2} \\ -2 & -1 \end{pmatrix}$. Hay que verificar la respuesta. Tenemos que

$$AA^{-1} = \begin{pmatrix} 2 & -3 \\ -4 & 5 \end{pmatrix}\begin{pmatrix} -\frac{5}{2} & -\frac{3}{2} \\ -2 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

y

$$A^{-1}A = \begin{pmatrix} -\frac{5}{2} & -\frac{3}{2} \\ -2 & -1 \end{pmatrix}\begin{pmatrix} 2 & -3 \\ -4 & 5 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

De modo que A es invertible y $A^{-1} = \begin{pmatrix} -\frac{5}{2} & -\frac{3}{2} \\ -2 & -1 \end{pmatrix}$.

Ejemplo 3 Si $A = \begin{pmatrix} 1 & 2 \\ -2 & -4 \end{pmatrix}$, calcular A^{-1}, si existe.

Solución Si $A^{-1} = \begin{pmatrix} x & y \\ z & w \end{pmatrix}$ existe, entonces

$$AA^{-1} = \begin{pmatrix} 1 & 2 \\ -2 & -4 \end{pmatrix}\begin{pmatrix} x & y \\ z & w \end{pmatrix} = \begin{pmatrix} x + 2z & y + 2w \\ -2x - 4z & -2y - 4w \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

lo cual conduce al sistema

$$\begin{array}{rrrrl} x & & +2z & & = 1 \\ & y & & +2w & = 0 \\ -2x & & -4z & & = 0 \\ & -2y & & -4w & = 1 \end{array} \qquad (10)$$

Usando el mismo razonamiento que en el Ejemplo 2, es posible escribir este sistema en la forma aumentada $(A\,|\,I)$ y reducir por renglones:

$$\left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ -2 & -4 & 0 & 1 \end{array}\right) \xrightarrow{A_{1,2}(2)} \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & 0 & 2 & 1 \end{array}\right)$$

No podemos continuar. La última fila dice que $0 = 2$ o que $0 = 1$, dependiendo de cual de los sistemas (el de x, z o el de y, w) se resuelve. Así pues, concluimos que el sistema (10) es incompatible, y por lo tanto, que A no es invertible.

Los últimos dos ejemplos ilustran un procedimiento que siempre da resultado cuando se trata de invertir una matriz.

Procedimiento para calcular la inversa de una matriz cuadrada A

Paso 1. Escribir la matriz aumentada $(A\,|\,I)$.

Paso 2. Utilizar reducción por renglones para reducir la matriz A a su forma escalonada.

Paso 3. Decidir si la matriz A es invertible.

(a) Si A puede ser reducida a la matriz identidad I, entonces A es la matriz a la derecha de la barra vertical.

(b) Si la reducción por renglones conduce a algún renglón de ceros a la izquierda de la barra vertical, la matriz A no es invertible.

Observación. Podemos expresar (a) y (b) como sigue

Una matriz cuadrada A es invertible, si y sólo si su forma escalonada, reducida por renglones, es la matriz identidad.

Sea $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$. Entonces, como en la Ecuación (1.2.7), definimos

$$\text{Determinante de } A = a_{11}a_{22} - a_{12}a_{21} \qquad (11)$$

Abreviamos el determinante de A como det A.

Teorema 4 Sea A una matriz de 2×2. Entonces:

i. A es invertible si y sólo si det $A \neq 0$.
ii. Si det $A \neq 0$, entonces

$$A^{-1} = \frac{1}{\det A}\begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}^{*} \qquad (12)$$

* Esta fórmula puede ser obtenida directamente aplicando el procedimiento para determinar una inversa (ver el Problema 46).

Demostración Primero supongamos que $\det A \neq 0$ y sea $B = (1/\det A)\begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$. Entonces

$$BA = \frac{1}{\det A}\begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

$$= \frac{1}{a_{11}a_{22} - a_{12}a_{21}}\begin{pmatrix} a_{22}a_{11} - a_{12}a_{21} & 0 \\ 0 & -a_{21}a_{12} + a_{11}a_{22} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I$$

De la misma forma $AB = I$, lo que muestra que A es invertible y que $B = A^{-1}$. Debemos aún mostrar que si A es invertible entonces $\det A \neq 0$. Para hacer esto consideremos el sistema

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1 \\ a_{21}x_1 + a_{22}x_2 &= b_2 \end{aligned} \tag{13}$$

Hacemos esto porque de acuerdo con el Teorema Resumen (Teorema 1.2.1), si este sistema tiene una única solución, entonces su determinante es distinto de cero. El sistema se puede escribir en la forma

$$A\mathbf{x} = \mathbf{b} \tag{14}$$

con $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ y $\mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$. Entonces, como A es invertible, vemos de (2) que el sistema (14) tiene una única solución dada por

$$\mathbf{x} = A^{-1}\mathbf{b}$$

Pero por el Teorema 1.2.1 el hecho de que el sistema (13) tenga una única solución implica que $\det A \neq 0$. Esto completa la demostración. ■

Ejemplo 4 Sea $A = \begin{pmatrix} 2 & -4 \\ 1 & 3 \end{pmatrix}$. Calcule A^{-1} si es que existe.

Solución Se tiene que $\det A = (2)(3) - (-4)(1) = 10$; por consiguiente A^{-1} existe. De la Ecuación (12), resulta

$$A^{-1} = \frac{1}{10}\begin{pmatrix} 3 & 4 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} \frac{3}{10} & \frac{4}{10} \\ -\frac{1}{10} & \frac{2}{10} \end{pmatrix}$$

Verificación

$$A^{-1}A = \frac{1}{10}\begin{pmatrix} 3 & 4 \\ -1 & 2 \end{pmatrix}\begin{pmatrix} 2 & -4 \\ 1 & 3 \end{pmatrix} = \frac{1}{10}\begin{pmatrix} 10 & 0 \\ 0 & 10 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

y

$$AA^{-1} = \begin{pmatrix} 2 & -4 \\ 1 & 3 \end{pmatrix}\begin{pmatrix} \frac{3}{10} & \frac{4}{10} \\ -\frac{1}{10} & \frac{2}{10} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Ejemplo 5 Sea $A = \begin{pmatrix} 1 & 2 \\ -2 & -4 \end{pmatrix}$. Calcule A^{-1} si es que existe.

Solución Encontramos que $\det A = (1)(-4) - (2)(-2) = -4 + 4 = 0$, de donde A^{-1} no existe, como vimos en el Ejemplo 3.

El procedimiento anterior funciona para matrices de $n \times n$ con $n > 2$. Ilustraremos esto con una serie de ejemplos.

Ejemplo 6 Sea $A = \begin{pmatrix} 2 & 4 & 6 \\ 4 & 5 & 6 \\ 3 & 1 & -2 \end{pmatrix}$ (vea Ejemplo 1.6.1). Calcule A^{-1} si es que existe.

Solución Primero ponemos I en seguida de A en forma de matriz aumentada

$$\left(\begin{array}{ccc|ccc} 2 & 4 & 6 & 1 & 0 & 0 \\ 4 & 5 & 6 & 0 & 1 & 0 \\ 3 & 1 & -2 & 0 & 0 & 1 \end{array}\right)$$

y llevamos a cabo la reducción por renglones.

$$\xrightarrow{M_1(\frac{1}{2})} \left(\begin{array}{ccc|ccc} 1 & 2 & 3 & \frac{1}{2} & 0 & 0 \\ 4 & 5 & 6 & 0 & 1 & 0 \\ 3 & 1 & -2 & 0 & 0 & 1 \end{array}\right) \xrightarrow[A_{1,3}(-3)]{A_{1,2}(-4)} \left(\begin{array}{ccc|ccc} 1 & 2 & 3 & \frac{1}{2} & 0 & 0 \\ 0 & -3 & -6 & -2 & 1 & 0 \\ 0 & -5 & -11 & -\frac{3}{2} & 0 & 1 \end{array}\right)$$

$$\xrightarrow{M_2(-\frac{1}{3})} \left(\begin{array}{ccc|ccc} 1 & 2 & 3 & \frac{1}{2} & 0 & 0 \\ 0 & 1 & 2 & \frac{2}{3} & -\frac{1}{3} & 0 \\ 0 & -5 & -11 & -\frac{3}{2} & 0 & 1 \end{array}\right) \xrightarrow[A_{2,3}(5)]{A_{2,1}(-2)} \left(\begin{array}{ccc|ccc} 1 & 0 & -1 & -\frac{5}{6} & \frac{2}{3} & 0 \\ 0 & 1 & 2 & \frac{2}{3} & -\frac{1}{3} & 0 \\ 0 & 0 & -1 & \frac{11}{6} & -\frac{5}{3} & 1 \end{array}\right)$$

$$\xrightarrow{M_3(-1)} \left(\begin{array}{ccc|ccc} 1 & 0 & -1 & -\frac{5}{6} & \frac{2}{3} & 0 \\ 0 & 1 & 2 & \frac{2}{3} & -\frac{1}{3} & 0 \\ 0 & 0 & 1 & -\frac{11}{6} & \frac{5}{3} & -1 \end{array}\right) \xrightarrow[A_{3,2}(-2)]{A_{3,1}(1)} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{8}{3} & \frac{7}{3} & -1 \\ 0 & 1 & 0 & \frac{13}{3} & -\frac{11}{3} & 2 \\ 0 & 0 & 1 & -\frac{11}{6} & \frac{5}{3} & -1 \end{array}\right)$$

Como ahora A ha sido reducida a I tenemos

$$A^{-1} = \begin{pmatrix} -\frac{8}{3} & \frac{7}{3} & -1 \\ \frac{13}{3} & -\frac{11}{3} & 2 \\ -\frac{11}{6} & \frac{5}{3} & -1 \end{pmatrix} = \frac{1}{6}\begin{pmatrix} -16 & 14 & -6 \\ 26 & -22 & 12 \\ -11 & 10 & -6 \end{pmatrix}$$

Se saca $\frac{1}{6}$ como factor común.

Verificación $A^{-1}A = \frac{1}{6}\begin{pmatrix} -16 & 14 & -6 \\ 26 & -22 & 12 \\ -11 & 10 & -6 \end{pmatrix}\begin{pmatrix} 2 & 4 & 6 \\ 4 & 5 & 6 \\ 3 & 1 & -2 \end{pmatrix} = \frac{1}{6}\begin{pmatrix} 6 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 6 \end{pmatrix} = I.$

También podemos verificar que $AA^{-1} = I$.

Advertencia. Es muy fácil cometer errores numéricos al calcular A^{-1}. Por tanto, es esencial verificar los cálculos probando que $A^{-1}A = I$.

Ejemplo 7 Sea $A = \begin{pmatrix} 2 & 4 & 3 \\ 0 & 1 & -1 \\ 3 & 5 & 7 \end{pmatrix}$. Calcule A^{-1} si es que existe.

Solución Procediendo como en el Ejemplo 6 obtenemos, sucesivamente, las siguientes matrices aumentadas:

$$\left(\begin{array}{ccc|ccc} 2 & 4 & 3 & 1 & 0 & 0 \\ 0 & 1 & -1 & 0 & 1 & 0 \\ 3 & 5 & 7 & 0 & 0 & 1 \end{array}\right) \xrightarrow{M_1(\frac{1}{2})} \left(\begin{array}{ccc|ccc} 1 & 2 & \frac{3}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 1 & -1 & 0 & 1 & 0 \\ 3 & 5 & 7 & 0 & 0 & 1 \end{array}\right)$$

$$\xrightarrow{A_{1,3}(-3)} \left(\begin{array}{ccc|ccc} 1 & 2 & \frac{3}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 1 & -1 & 0 & 1 & 0 \\ 0 & -1 & \frac{5}{2} & -\frac{3}{2} & 0 & 1 \end{array}\right) \xrightarrow[A_{2,3}(1)]{A_{2,1}(-2)} \left(\begin{array}{ccc|ccc} 1 & 0 & \frac{7}{2} & \frac{1}{2} & -2 & 0 \\ 0 & 1 & -1 & 0 & 1 & 0 \\ 0 & 0 & \frac{3}{2} & -\frac{3}{2} & 1 & 1 \end{array}\right)$$

$$\xrightarrow{M_3(\frac{2}{3})} \left(\begin{array}{ccc|ccc} 1 & 0 & \frac{7}{2} & \frac{1}{2} & -2 & 0 \\ 0 & 1 & -1 & 0 & 1 & 0 \\ 0 & 0 & 1 & -1 & \frac{2}{3} & \frac{2}{3} \end{array}\right) \xrightarrow[A_{3,2}(1)]{A_{3,1}(-\frac{7}{2})} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 4 & -\frac{13}{3} & -\frac{7}{3} \\ 0 & 1 & 0 & -1 & \frac{5}{3} & \frac{2}{3} \\ 0 & 0 & 1 & -1 & \frac{2}{3} & \frac{2}{3} \end{array}\right)$$

Por consiguiente

$$A^{-1} = \begin{pmatrix} 4 & -\frac{13}{3} & -\frac{7}{3} \\ -1 & \frac{5}{3} & \frac{2}{3} \\ -1 & \frac{2}{3} & \frac{2}{3} \end{pmatrix}$$

Verificación $$A^{-1}A = \begin{pmatrix} 4 & -\frac{13}{3} & -\frac{7}{3} \\ -1 & \frac{5}{3} & \frac{2}{3} \\ -1 & \frac{2}{3} & \frac{2}{3} \end{pmatrix} \begin{pmatrix} 2 & 4 & 3 \\ 0 & 1 & -1 \\ 3 & 5 & 7 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Ejemplo 8 Sea $A = \begin{pmatrix} 1 & -3 & 4 \\ 2 & -5 & 7 \\ 0 & -1 & 1 \end{pmatrix}$. Calcule A^{-1} si es que existe.

Solución Procediendo como antes obtenemos, sucesivamente,

$$\left(\begin{array}{ccc|ccc} 1 & -3 & 4 & 1 & 0 & 0 \\ 2 & -5 & 7 & 0 & 1 & 0 \\ 0 & -1 & 1 & 0 & 0 & 1 \end{array}\right) \xrightarrow{A_{1,2}(-2)} \left(\begin{array}{ccc|ccc} 1 & -3 & 4 & 1 & 0 & 0 \\ 0 & 1 & -1 & -2 & 1 & 0 \\ 0 & -1 & 1 & 0 & 0 & 1 \end{array}\right)$$

$$\xrightarrow[A_{2,3}(1)]{A_{2,1}(3)} \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & -5 & 3 & 0 \\ 0 & 1 & -1 & -2 & 1 & 0 \\ 0 & 0 & 0 & -2 & 1 & 1 \end{array}\right)$$

Esto es lo más que podemos hacer. La matriz A *no se puede* reducir a la matriz identidad y concluimos que A *no* es invertible.

Hay otra manera de ver el resultado del último ejemplo. Sea **b** cualquier vector de tres componentes y consideremos el sistema $A\mathbf{x} = \mathbf{b}$. Si tratásemos de resolver este sistema por eliminación gaussiana, se terminaría con una ecuación de la forma $0 = c \neq 0$ como en el Ejemplo 3, o también con $0 = 0$. Este es el caso (*ii*) o (*iii*) de la Sección 1.6 (ver la página 41). Esto es, el sistema no tiene solución o posee un número infinito de soluciones. La única posibilidad que se excluye es la de que el sistema tenga una solución única. Pero si A^{-1} existiera, entonces habría una sola solución dada por $\mathbf{x} = A^{-1}\mathbf{b}$. Concluimos pues que

> Si en la reducción por filas de A se termina con una fila de ceros, la matriz A no es invertible.

Definición 3 ***Matrices equivalentes por renglones (o filas).*** Supongamos que por medio de operaciones elementales sobre filas, podemos transformar una matriz A en una matriz B. Entonces se dice que A y B son *equivalentes por renglones*.

El razonamiento usado anteriormente se puede utilizar para la demostración del siguiente teorema (ver el Problema 47).

Teorema 5 Sea A una matriz de $n \times n$.

i. A es invertible si y sólo si A es equivalente por renglones a la matriz identidad I_n; esto es, la forma en escalón por filas reducida de A es I_n.
ii. A es invertible si y sólo si el sistema $A\mathbf{x} = \mathbf{b}$ tiene solución única para todo vector **b** de n componentes.
iii. Si A es invertible, esta solución única está dada por $\mathbf{x} = A^{-1}\mathbf{b}$.

Ejemplo 9 Resolver el sistema

$$\begin{aligned} 2x_1 + 4x_2 + 3x_3 &= 6 \\ x_2 - x_3 &= -4 \\ 3x_1 + 5x_2 + 7x_3 &= 7 \end{aligned}$$

Solución Este sistema puede ser expresado como $A\mathbf{x} = \mathbf{b}$, en donde $A = \begin{pmatrix} 2 & 4 & 3 \\ 0 & 1 & -1 \\ 3 & 5 & 7 \end{pmatrix}$ y $\mathbf{b} = \begin{pmatrix} 6 \\ -4 \\ 7 \end{pmatrix}$. En el Ejemplo 7 encontramos que A^{-1} existe y que

$$A^{-1} = \begin{pmatrix} 4 & -\frac{13}{3} & -\frac{7}{3} \\ -1 & \frac{5}{3} & \frac{2}{3} \\ -1 & \frac{2}{3} & \frac{2}{3} \end{pmatrix}$$

Así pues, la solución única está dada por

$$\mathbf{x}=\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=A^{-1}\mathbf{b}=\begin{pmatrix}4&-\frac{13}{3}&-\frac{7}{3}\\-1&\frac{5}{3}&\frac{2}{3}\\-1&\frac{2}{3}&\frac{2}{3}\end{pmatrix}\begin{pmatrix}6\\-4\\7\end{pmatrix}=\begin{pmatrix}25\\-8\\-4\end{pmatrix}$$

Ejemplo 10 En el modelo de insumo-producto de Leontief descrito en el Ejemplo 1.6.8, se obtuvo el sistema

$$\begin{aligned}a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n + e_1 &= x_1\\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n + e_2 &= x_2\\ \vdots \qquad \vdots \qquad\quad \vdots \qquad \vdots \quad &\ \ \vdots\\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n + e_n &= x_n\end{aligned} \qquad (15)$$

que puede ser escrito como

$$A\mathbf{x} + \mathbf{e} = \mathbf{x} = I\mathbf{x}$$

o también,

$$(I - A)\mathbf{x} = \mathbf{e} \qquad (16)$$

La matriz A de demandas internas lleva el nombre de *matriz de tecnología*, y la matriz $I - A$ se llama *matriz de Leontief*. Si la matriz de Leontief es invertible, entonces los sistemas (15) y (16) tienen soluciones únicas.

Leontief usó su modelo para analizar la economía estadunidense de 1958.* Dividió la economía en 81 sectores y los agrupó en seis familias de sectores relacionados. Para simplificar, se trata cada familia de sectores como uno solo, de modo que se manejará la economía de Estados Unidos como una de seis industrias principales. Éstas aparecen enumeradas en la Tabla 1.1.

Tabla 1.1

Sector (industrial)	Ejemplos
Final no metálica (FN)	Muebles, alimentos procesados
Final metálica (FM)	Enseres domésticos, vehículos de motor
Básica metálica (BM)	Productos elaborados con maquinaria, minería
Básica no metálica (BN)	Agricultura, artes gráficas
Energía (E)	Petróleo, carbón
Servicios (S)	Diversiones, bienes raíces

El cuadro de insumo-producto, Tabla 1.2, proporciona las demandas internas en 1958 con base en las cifras de Leontief. Las unidades en la tabla son millones de dólares ($). Así, por ejemplo, el número 0.173 en la posición 6,5 significa que es necesario suministrar 0.173 millones = $173,000 de servicios para producir $1,000,000 equivalente de energía. Similarmente, el número 0.037 en la posición 4,2 significa que para producir el equivalente a $1,000,000 de productos finales metálicos, es necesario gastar $37,000 en productos básicos no metálicos.

* *Scientific American* (abril de 1965), págs. 26-27.

Tabla 1.2 **Demandas internas en la economía de Estados Unidos en 1958**

	FN	FM	BM	BN	E	S
FN	0.170	0.004	0	0.029	0	0.008
FM	0.003	0.295	0.018	0.002	0.004	0.016
BM	0.025	0.173	0.460	0.007	0.011	0.007
BN	0.348	0.037	0.021	0.403	0.011	0.048
E	0.007	0.001	0.039	0.025	0.358	0.025
S	0.120	0.074	0.104	0.123	0.173	0.234

Finalmente, Leontief estimó las siguientes demandas externas en el año de 1958 para la economía estadunidense (en millones de dólares).

Tabla 1.3 **Demandas externas en 1958 para la economía de E.U. (millones de dólares)**

FN	$99,640
FM	$75,548
BM	$14,444
BN	$33,501
E	$23,527
S	$263,985

¿Cuántas unidades de cada uno de los seis sectores es necesario producir para activar la economía estadunidense de 1958 a fin de poder satisfacer todas las demandas externas?

Solución La matriz de tecnología está dada por

$$A=\begin{pmatrix}0.170&0.004&0&0.029&0&0.008\\0.003&0.295&0.018&0.002&0.004&0.016\\0.025&0.173&0.460&0.007&0.011&0.007\\0.348&0.037&0.021&0.403&0.011&0.048\\0.007&0.001&0.039&0.025&0.358&0.025\\0.120&0.074&0.104&0.123&0.173&0.234\end{pmatrix}$$

y

$$\mathbf{e}=\begin{pmatrix}99,640\\75,548\\14,444\\33,501\\23,527\\263,985\end{pmatrix}$$

Para obtener la matriz de Leontief, obtenemos, restando,

$$I - A = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} - \begin{pmatrix} 0.170 & 0.004 & 0 & 0.029 & 0 & 0.008 \\ 0.003 & 0.295 & 0.018 & 0.002 & 0.004 & 0.016 \\ 0.025 & 0.173 & 0.460 & 0.007 & 0.011 & 0.007 \\ 0.348 & 0.037 & 0.021 & 0.403 & 0.011 & 0.048 \\ 0.007 & 0.001 & 0.039 & 0.025 & 0.358 & 0.025 \\ 0.120 & 0.074 & 0.104 & 0.123 & 0.173 & 0.234 \end{pmatrix}$$

$$= \begin{pmatrix} 0.830 & -0.004 & 0 & -0.029 & 0 & -0.008 \\ -0.003 & 0.705 & -0.018 & -0.002 & -0.004 & -0.016 \\ -0.025 & -0.173 & 0.540 & -0.007 & -0.011 & -0.007 \\ -0.348 & -0.037 & -0.021 & 0.597 & -0.011 & -0.048 \\ -0.007 & -0.001 & -0.039 & -0.025 & 0.642 & -0.025 \\ -0.120 & -0.074 & -0.104 & -0.123 & -0.173 & 0.766 \end{pmatrix}$$

La determinación de la inversa de una matriz de 6 × 6 es un asunto tedioso. Usando tres decimales en una calculadora, se obtiene la matriz siguiente. Los pasos intermedios se omiten.

$$(I - A)^{-1} = \begin{pmatrix} 1.234 & 0.014 & 0.006 & 0.064 & 0.007 & 0.018 \\ 0.017 & 1.436 & 0.057 & 0.012 & 0.020 & 0.032 \\ 0.071 & 0.465 & 1.877 & 0.019 & 0.045 & 0.031 \\ 0.751 & 0.134 & 0.100 & 1.740 & 0.066 & 0.124 \\ 0.060 & 0.045 & 0.130 & 0.082 & 1.578 & 0.059 \\ 0.339 & 0.236 & 0.307 & 0.312 & 0.376 & 1.349 \end{pmatrix}$$

Así pues, el vector "ideal" de producción (o producto) está dado por

$$\mathbf{x} = (I - A)^{-1}\mathbf{e} = \begin{pmatrix} 1.234 & 0.014 & 0.006 & 0.064 & 0.007 & 0.018 \\ 0.017 & 1.436 & 0.057 & 0.012 & 0.020 & 0.032 \\ 0.071 & 0.465 & 1.877 & 0.019 & 0.045 & 0.031 \\ 0.751 & 0.134 & 0.100 & 1.740 & 0.066 & 0.124 \\ 0.060 & 0.045 & 0.130 & 0.082 & 1.578 & 0.059 \\ 0.339 & 0.236 & 0.307 & 0.312 & 0.376 & 1.349 \end{pmatrix} \begin{pmatrix} 99{,}640 \\ 75{,}548 \\ 14{,}444 \\ 33{,}501 \\ 23{,}527 \\ 263{,}985 \end{pmatrix}$$

$$= \begin{pmatrix} 131{,}161 \\ 120{,}324 \\ 79{,}194 \\ 178{,}936 \\ 66{,}703 \\ 426{,}542 \end{pmatrix}$$

Esto significa que se requieren 131,161 unidades (con valor de $131,161 millones) de productos finales no metálicos; 120,324 unidades de productos metálicos finales; 79,194 unidades de productos metálicos básicos; 178,936 unidades de productos básicos no metálicos; 66,703 unidades de energía, y 426,542 unidades de servicio para activar la economía y satisfacer las demandas externas de 1958.

En la Sección 1.2 encontramos la primera forma del Teorema Resumen (Teorema 1.2.1). Ahora estamos en condiciones de mejorarlo. El siguiente teorema establece la equivalencia de varias afirmaciones referentes a los conceptos de inversa, unicidad de soluciones, equivalencia por renglón y determinantes. Esto significa que si una de las afirmaciones es verdadera, todas las demás también lo son. En este punto podemos probar la equivalencia de las partes (*i*), (*ii*), (*iii*) y (*iv*). Terminaremos la demostración después de haber desarrollado una teoría básica sobre determinantes (vea Teorema 2.4.4).

Teorema 6 ***Teorema Resumen, Versión 2.*** Sea A una matriz de $n \times n$. Entonces cada una de las cinco siguientes afirmaciones implica a las otras cuatro (esto es, si una es cierta, todas son ciertas):

i. A es invertible.
ii. La única solución al sistema homogéneo $A\mathbf{x}=\mathbf{0}$ es la solución trivial ($\mathbf{x}=\mathbf{0}$).
iii. El sistema $A\mathbf{x}=\mathbf{b}$ tiene una única solución para cada n-vector $\mathbf{b}$.
iv. A es equivalente por renglones a la matriz identidad de $n \times n$ I_n.
v. det $A \neq 0$ (Hasta ahora, det A sólo está definido si A es una matriz de 2 × 2.)

Demostración Ya hemos visto que las afirmaciones (*i*) y (*iii*) son equivalentes (Teorema 5 (parte *ii*)) y que (*i*) y (*iv*) son equivalentes (Teorema 5 (parte *i*)). Veremos que (*ii*) y (*iv*) son equivalentes. Supongamos que (*ii*) es cierta. Esto es, supongamos que $A\mathbf{x} = \mathbf{0}$ tiene solamente la solución trivial $\mathbf{x} = \mathbf{0}$. Si escribimos este sistema obtenemos

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0 \\ \vdots \qquad \vdots \qquad\quad & \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= 0 \end{aligned} \qquad (17)$$

Si A no fuera equivalente a I_n entonces la reducción por renglones de la matriz aumentada asociada a (17) nos llevaría a un renglón de ceros. Pero si, por ejemplo, el último renglón es cero, entonces la última ecuación quedaría 0 = 0. Entonces el sistema homogéneo se reduce a un sistema de $n - 1$ ecuaciones con n incógnitas, el cual, en virtud del Teorema 1.7.1 tiene un número infinito de soluciones. Pero supusimos que $\mathbf{x} = \mathbf{0}$ era la única solución al sistema (17). Esta contradicción muestra que A es equivalente por renglones a I_n. Inversa-

mente, supongamos que (*iv*) es cierta; esto es, supongamos que A es equivalente por renglones a I_n. Entonces por el Teorema 5 (parte *i*), A es invertible y por el Teorema 5 (parte *iii*) la única solución a $A\mathbf{x} = \mathbf{0}$ es $\mathbf{x} = A^{-1}\mathbf{0} = \mathbf{0}$. Así, (*ii*) y (*iv*) son equivalentes. En el Teorema 1.2.1 mostramos que (*i*) y (*vi*) son equivalentes en el caso 2×2. Probaremos la equivalencia de (*ii*), (*v*) y (*vi*) en la Sección 2.4. ■

Observación. Podríamos añadir otra afirmación al teorema. Supongamos que el sistema $A\mathbf{x} = \mathbf{b}$ tiene una única solución. Sea R una matriz en forma escalonada que sea equivalente por renglones a A. Entonces R no puede tener un renglón de ceros, pues si lo tuviera no podría ser reducida a la matriz identidad.* Así, la forma escalonada de A debería parecerse a lo siguiente:

$$\begin{pmatrix} 1 & r_{12} & r_{13} & \cdots & r_{1n} \\ 0 & 1 & r_{23} & \cdots & r_{2n} \\ 0 & 0 & 1 & \cdots & r_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix} \qquad (18)$$

Esto es, R es una matriz con unos en la diagonal y ceros abajo de ella. Tenemos así el Teorema 7.

Teorema 7 Si se da alguna de las condiciones del Teorema 6, entonces la forma escalonada de A tiene la forma de la matriz (18).

Hemos visto que para verificar que $B = A^{-1}$ debemos comprobar que $AB = BA = I$. Veremos que basta con probar la mitad de esto.

Teorema 8 Sean A y B matrices de $n \times n$. Entonces A es invertible y $B = A^{-1}$ si (*i*) $BA = I$ o bien (*ii*) $AB = I$.

Observación. Este teorema simplifica el trabajo de ver si una matriz es la inversa de otra.

Demostración

i. Supongamos que $BA = I$. Consideremos el sistema homogéneo $A\mathbf{x} = \mathbf{0}$. Multiplicando por la izquierda ambos lados de esta ecuación por B obtenemos

$$BA\mathbf{x} = B\,\mathbf{0} \qquad (19)$$

Pero $BA = I$ y $B\mathbf{0} = \mathbf{0}$, de donde (19) se convierte en $I\mathbf{x} = \mathbf{0}$ o $\mathbf{x} = \mathbf{0}$. Esto muestra que $\mathbf{x} = \mathbf{0}$ es la única solución a $A\mathbf{x} = \mathbf{0}$ y, por el Teorema 6 (parte *i* y *ii*), esto significa que A es invertible. Debemos mostrar aún

* Notemos que si el renglón i-ésimo de R contiene solamente ceros, entonces el sistema homogéneo $R\mathbf{x} = \mathbf{0}$ contiene más incógnitas que ecuaciones (pues la i-ésima ecuación es la ecuación cero) y el sistema tiene un número infinito de soluciones. Pero entonces $A\mathbf{x} = \mathbf{0}$ tiene un número infinito de soluciones, lo que contradice nuestra suposición.

que $B = A^{-1}$. Sea $A^{-1} = C$. Entonces $AC = I$. Así, $BAC = B(AC) = BI = B$ y $BAC = (BA)C = IC = C$. Por tanto $B = C$ y la parte (*i*) queda demostrada.

ii. Sea $AB = I$. Entonces, de la parte (*i*), $A = B^{-1}$. De la Definición 2 esto significa que $AB = BA = I$, lo que prueba que A es invertible y que $B = A^{-1}$. Esto completa la demostración. ■

Problemas 1.8

En los Problemas del 1 *al* 15 *determine si la matriz dada es invertible. Si lo es, calcule su inversa.*

1. $\begin{pmatrix} 2 & 1 \\ 3 & 2 \end{pmatrix}$
2. $\begin{pmatrix} -1 & 6 \\ 2 & -12 \end{pmatrix}$
3. $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$
4. $\begin{pmatrix} 1 & 1 \\ 3 & 3 \end{pmatrix}$
5. $\begin{pmatrix} a & a \\ b & b \end{pmatrix}$
6. $\begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 3 \\ 5 & 5 & 1 \end{pmatrix}$
7. $\begin{pmatrix} 3 & 2 & 1 \\ 0 & 2 & 2 \\ 0 & 0 & -1 \end{pmatrix}$
8. $\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$
9. $\begin{pmatrix} 1 & 6 & 2 \\ -2 & 3 & 5 \\ 7 & 12 & -4 \end{pmatrix}$
10. $\begin{pmatrix} 3 & 1 & 0 \\ 1 & -1 & 2 \\ 1 & 1 & 1 \end{pmatrix}$
11. $\begin{pmatrix} 2 & -1 & 4 \\ -1 & 0 & 5 \\ 19 & -7 & 3 \end{pmatrix}$
12. $\begin{pmatrix} 1 & 2 & 3 \\ 1 & 1 & 2 \\ 0 & 1 & 2 \end{pmatrix}$
13. $\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & -1 & 2 \\ 1 & -1 & 2 & 1 \\ 1 & 3 & 3 & 2 \end{pmatrix}$
14. $\begin{pmatrix} 1 & 0 & 2 & 3 \\ -1 & 1 & 0 & 4 \\ 2 & 1 & -1 & 3 \\ -1 & 0 & 5 & 7 \end{pmatrix}$
15. $\begin{pmatrix} 1 & -3 & 0 & -2 \\ 3 & -12 & -2 & -6 \\ -2 & 10 & 2 & 5 \\ -1 & 6 & 1 & 3 \end{pmatrix}$
16. Muestre que si A, B y C son matrices invertibles entonces ABC es invertible y $(ABC)^{-1} = C^{-1}B^{-1}A^{-1}$.
17. Si $A_1, A_2, \ldots, A_n$ son matrices invertibles de $n \times n$ muestre que $A_1A_2 \cdots A_n$ es invertible y calcule su inversa.
18. Muestre que la matriz $\begin{pmatrix} 3 & 4 \\ -2 & -3 \end{pmatrix}$ es igual a su propia inversa.
19. Muestre que la matriz $\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ es igual a su propia inversa si $A = \pm I$ o si $a_{11} = -a_{22}$ y $a_{21}a_{12} = 1 - a_{11}^2$.
20. Encuentre el vector de salidas $\mathbf{x}$ en el modelo de insumo-producto de Leontief si
$$n = 3,\ \mathbf{e} = \begin{pmatrix} 30 \\ 20 \\ 40 \end{pmatrix} \quad \text{y} \quad A = \begin{pmatrix} \frac{1}{5} & \frac{1}{5} & 0 \\ \frac{2}{5} & \frac{2}{5} & \frac{3}{5} \\ \frac{1}{5} & \frac{1}{10} & \frac{2}{5} \end{pmatrix}.$$
21. ★ Suponga que A es de $n \times m$ y B es de $m \times n$ de forma que AB es de $n \times n$. Muestre que AB no es invertible si $n > m$. [*Sugerencia:* Muestre que existe un vector $\mathbf{x}$ diferente de cero tal que $AB\mathbf{x} = \mathbf{0}$ y luego aplique el Teorema 6.]
22. ★ Utilice los métodos de esta sección para encontrar las inversas de las siguientes matrices con componentes complejas:

a. $\begin{pmatrix} i & 2 \\ 1 & -i \end{pmatrix}$ b. $\begin{pmatrix} 1-i & 0 \\ 0 & 1+i \end{pmatrix}$ c. $\begin{pmatrix} 1 & i & 0 \\ -i & 0 & 1 \\ 0 & 1+i & 1-i \end{pmatrix}$

23. Muestre que para cualquier número real θ la matriz $\begin{pmatrix} \operatorname{sen}\theta & \cos\theta & 0 \\ \cos\theta & -\operatorname{sen}\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$ es invertible y encuentre su inversa.

24. Calcule la inversa de $A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{pmatrix}$.

25. Una matriz cuadrada $A = (a_{ij})$ se conoce como *diagonal* si todos sus elementos que no están en la diagonal principal son cero. Esto es, $a_{ij} = 0$ si $i \neq j$. (La matriz del Problema 24 es diagonal.) Muestre que una matriz diagonal es invertible si y sólo si cada una de sus componentes diagonales es diferente de cero.

26. Sea

$$A = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}$$

una matriz diagonal tal que cada una de sus componentes diagonales es distinta de cero. Calcule A^{-1}.

27. Calcule la inversa de $A = \begin{pmatrix} 2 & 1 & -1 \\ 0 & 3 & 4 \\ 0 & 0 & 5 \end{pmatrix}$.

28. Muestre que la matriz $A = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 0 & 0 \\ 4 & 6 & 1 \end{pmatrix}$ no es invertible.

★ 29. Una matriz cuadrada se conoce como *triangular superior* (*inferior*) si todos sus elementos por debajo (por arriba) de la diagonal principal son cero. (La matriz del Problema 27 es triangular superior y la matriz del Problema 28 es triangular inferior.) Muestre que una matriz triangular superior o inferior es invertible si y sólo si cada uno de sus elementos diagonales es diferente de cero.

★ 30. Muestre que la inversa de una matriz triangular superior invertible es triangular superior. [*Sugerencia:* Demuestre primero el resultado para una matriz de 3 × 3.]

En los Problemas 31 *y* 32 *se da una matriz. En cada caso, demuestre que la matriz no es invertible encontrando un vector* **x** *distinto de cero tal que* $A\mathbf{x} = \mathbf{0}$.

31. $\begin{pmatrix} 2 & -1 \\ -4 & 2 \end{pmatrix}$

32. $\begin{pmatrix} 1 & -1 & 3 \\ 0 & 4 & -2 \\ 2 & -6 & 8 \end{pmatrix}$

33. Una fábrica de muebles finos tiene dos divisiones: un taller de máquinas en donde se fabrican las piezas de los muebles, y una división de ensamblaje y terminado en la cual las piezas son unidas y armadas en un producto terminado. Supóngase que hay 12 empleados en el taller y 20 en la división de terminado, y que cada empleado trabaja 8 horas al día. Supóngase, además, que la fábrica sólo produce dos tipos de muebles: sillas y mesas. Una silla requiere de 384/17 horas de máquina y 480/17 horas de ensamblaje y terminado. Una mesa requiere de 240/17 horas de máquina y 640/17 horas de ensamblaje y terminado. Suponiendo que existe una demanda ilimitada de estos productos y que el fabricante desea mantener a todos sus empleados ocupados, ¿cuántas sillas y mesas puede producir esta fábrica en un día?

34. El armario mágico de una bruja contiene 10 onzas de tréboles de cuatro hojas (molidos) y 14 onzas de raíz de mandrágora en polvo. El contenido del armario se restituye en forma automática siempre y cuando la hechicera utilice todos sus suministros. Un lote de poción de amor requiere de $3\frac{1}{13}$ onzas de tréboles y $2\frac{2}{13}$ onzas de mandrágora. Una receta de una cura bien conocida (entre los hechiceros) para el resfriado común requiere de $5\frac{5}{13}$ onzas de los tréboles y $10\frac{10}{13}$ onzas de la mandrágora. ¿Cuánto de la poción de amor y de la cura debe producir la bruja para exactamente vaciar el contenido de su armario?

35. Un campesino alimenta su ganado con una mezcla de dos tipos de alimento. Una unidad estándar del tipo A suministra a una cabeza de ganado 10% de sus requerimientos diarios mínimos de proteína y 15% de los carbohidratos. El tipo B contiene, en una unidad estándar, 12% del requerimiento de proteína y 8% del de carbohidratos. Si el campesino desea dar a sus animales el 100% de sus requerimientos mínimos, ¿cuántas unidades de alimento debe dar a cada cabeza de ganado diariamente?

36. Una versión harto simplista de una tabla de insumo-producto para la economía israelí en 1958 divide a la economía en tres sectores: agricultura, manufacturas y energía, con el resultado siguiente.*

	Agricultura	**Manufacturas**	**Energía**
Agricultura	0.293	0	0
Manufacturas	0.014	0.207	0.017
Energía	0.044	0.010	0.216

(a) ¿Cuántas unidades de producción del sector agrícola son necesarias para obtener una unidad de producto agrícola?
(b) ¿Cuántas unidades de producción en agricultura son necesarias para producir 200,000 unidades de producto agrícola?
(c) ¿Cuántas unidades de producción agrícola se necesitan para la producción de 50,000 unidades de energía?
(d) ¿Cuántas unidades de energía se necesitan para producir 50,000 unidades de productos agrícolas?

37. Continuando con el Problema 36, las exportaciones (en miles de libras israelíes) en 1958 fueron

Agricultura	13,213
Manufacturas	17,597
Energía	1,786

* Wassily Leontief, *Input-Output Economics* (Nueva York: Oxford University Press, 1966), 54-57.

(a) Calcule las matrices de tecnología y de Leontief.
(b) Determine el número de libras israelíes necesarias en cada uno de los sectores económicos para poder activar la economía de este modelo y dé el valor de cada producto en las exportaciones.

En los Problemas del 38 *al* 45 *encuentre la forma escalonada de la matriz dada y utilícela para determinar directamente si la matriz dada es invertible.*

38. La matriz del Problema 3.
39. La matriz del Problema 1.
40. La matriz del Problema 4.
41. La matriz del Problema 7.
42. La matriz del problema 9.
43. La matriz del Problema 11.
44. La matriz del Problema 13.
45. La matriz del Problema 14.

46. Sea $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ y suponga que $a_{11}a_{22} - a_{12}a_{21} \neq 0$. Deduzca la fórmula (12) reduciendo por renglones la matriz aumentada $\left(\begin{array}{cc|cc} a_{11} & a_{12} & 1 & 0 \\ a_{21} & a_{22} & 0 & 1 \end{array}\right)$.

47. Demuestre las partes (i) y (ii) del Teorema 5.

1.9 Transpuesta de una matriz

Cada matriz tiene una matriz correspondiente que, como veremos en el Capítulo 2, tiene propiedades similares a las de la matriz original.

Definición 1 **Transpuesta.** Sea $A = (a_{ij})$ una matriz de $m \times n$. Entonces la *transpuesta de* A, escrita A^t, es la matriz de $n \times m$ obtenida intercambiando los renglones y columnas de A. Abreviadamente podemos escribir $A^t = (a_{ji})$. En otras palabras,

$$\text{si } A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \text{ entonces } A^t = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{pmatrix} \quad (1)$$

Dicho simplemente, el i-ésimo renglón de A es la i-ésima columna de A^t y la j-ésima columna de A es el j-ésimo renglón de A^t.

Ejemplo 1 Encuentre las transpuestas de las matrices

$$A = \begin{pmatrix} 2 & 3 \\ 1 & 4 \end{pmatrix} \quad B = \begin{pmatrix} 2 & 3 & 1 \\ -1 & 4 & 6 \end{pmatrix} \quad C = \begin{pmatrix} 1 & 2 & -6 \\ 2 & -3 & 4 \\ 0 & 1 & 2 \\ 2 & -1 & 5 \end{pmatrix}$$

Solución Intercambiando los renglones y columnas de cada matriz obtenemos

$$A^t = \begin{pmatrix} 2 & 1 \\ 3 & 4 \end{pmatrix} \quad B^t = \begin{pmatrix} 2 & -1 \\ 3 & 4 \\ 1 & 6 \end{pmatrix} \quad C^t = \begin{pmatrix} 1 & 2 & 0 & 2 \\ 2 & -3 & 1 & -1 \\ -6 & 4 & 2 & 5 \end{pmatrix}$$

Notemos, por ejemplo, que 4 es la componente del renglón 2 y la columna 3 de C mientras que 4 es la componente del renglón 3 y la columna 2 de C^t. Esto es, el elemento 23 de C es el elemento 32 de C^t.

Teorema 1 Supongamos que $A = (a_{ij})$ es una matriz de $n \times m$ y $B = (b_{ij})$ es una matriz de $m \times p$. Entonces:

i. $(A^t)^t = A$ (2)
ii. $(AB)^t = B^tA^t$ (3)
iii. Si A y B son de $n \times m$, entonces $(A + B)^t = A^t + B^t$ (4)
iv. Si A es invertible, entonces A^t es invertible y $(A^t)^{-1} = (A^{-1})^t$ (5)

Demostración

i. Esto se dice directamente de la definición de transpuesta.

ii. Primero notemos que AB es una matriz de $n \times p$ y así $(AB)^t$ es de $p \times n$. También, B^t es de $p \times m$ y A^t es de $m \times n$, de donde B^tA^t es de $p \times n$. Así, ambas matrices en la Ecuación (3) son del mismo tamaño. Ahora, el ij-ésimo elemento de AB es $\sum_{k=1}^{m} a_{ik}b_{kj}$ y éste es el ji-ésimo elemento de $(AB)^t$. Sea $C = B^t$ y $D = A^t$. Entonces el ij-ésimo elemento c_{ij} de C es b_{ji} y el ij-ésimo elemento d_{ij} de D es a_{ji}. Así, el ji-ésimo elemento de CD es igual al ji-ésimo elemento de $B^tA^t = \sum_{k=1}^{m} c_{jk}d_{ki} = \sum_{k=1}^{m} b_{kj}a_{ik} = \sum_{k=1}^{m} a_{ik}b_{kj}$ es igual al ji-ésimo elemento de $(AB)^t$. Esto completa la demostración de la parte (ii).

iii. Esta parte se deja como ejercicio (Problema 11).

iv. Esta parte es más difícil. Se demuestra en la siguiente sección, en la página 80.

La transpuesta juega un papel importante en la teoría de matrices. Veremos en capítulos subsecuentes que A y A^t tienen varias propiedades en común. Como las columnas de A^t son los renglones de A seremos capaces de usar propiedades relacionadas con la transpuesta para concluir que lo que es cierto para los renglones de una matriz es cierto para sus columnas. Concluimos esta sección con una definición importante.

Definición 2 **Matriz simétrica.** Se dice que la matriz A de $n \times n$ (cuadrada) es *simétrica* si $A^t = A$.

Ejemplo 2 Las cuatro matrices siguientes son simétricas:

$$I \quad A = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} \quad B = \begin{pmatrix} 1 & -4 & 2 \\ -4 & 7 & 5 \\ 2 & 5 & 0 \end{pmatrix} \quad C = \begin{pmatrix} -1 & 2 & 4 & 6 \\ 2 & 7 & 3 & 5 \\ 4 & 3 & 8 & 0 \\ 6 & 5 & 0 & -4 \end{pmatrix}$$

Veremos la importancia de las matrices simétricas en los Capítulos 5 y 6.

Problemas 1.9

En los Problemas del 1 *al* 10 *encuentre la transpuesta de la matriz dada.*

1. $\begin{pmatrix} -1 & 4 \\ 6 & 5 \end{pmatrix}$ **2.** $\begin{pmatrix} 3 & 0 \\ 1 & 2 \end{pmatrix}$ **3.** $\begin{pmatrix} 2 & 3 \\ -1 & 2 \\ 1 & 4 \end{pmatrix}$ **4.** $\begin{pmatrix} 2 & -1 & 0 \\ 1 & 5 & 6 \end{pmatrix}$

5. $\begin{pmatrix} 1 & 2 & 3 \\ -1 & 0 & 4 \\ 1 & 5 & 5 \end{pmatrix}$ **6.** $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & -5 \\ 3 & -5 & 7 \end{pmatrix}$ **7.** $\begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}$ **8.** $\begin{pmatrix} 2 & -1 \\ 2 & 4 \\ 1 & 6 \\ 1 & 5 \end{pmatrix}$

9. $\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & j \end{pmatrix}$ **10.** $\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$

11. Sean A y B matrices de $n \times m$. Muestre, usando la Definición 1, que $(A + B)^t = A^t + B^t$.

12. Encuentre números α y β tales que $\begin{pmatrix} 2 & \alpha & 3 \\ 5 & -6 & 2 \\ \beta & 2 & 4 \end{pmatrix}$ sea simétrica.

13. Si A y B son matrices simétricas de $n \times n$ demuestre que $A + B$ es simétrica.

14. Si A y B son matrices simétricas de $n \times n$, muestre que $(AB)^t = BA$.

15. Para cualquier matriz A muestre que el producto AA^t está definido y es una matriz simétrica.

16. Muestre que cualquier matriz diagonal es simétrica (Problema 1.8.25).

17. Muestre que la transpuesta de cualquier matriz triangular superior es triangular inferior (Problema 1.8.29).

18. Se dice que una matriz cuadrada es *antisimétrica* si $A^t = -A$ (esto es, $a_{ij} = -a_{ji}$). ¿Cuáles de las siguientes matrices son antisimétricas?

a. $\begin{pmatrix} 1 & -6 \\ 6 & 0 \end{pmatrix}$ **b.** $\begin{pmatrix} 0 & -6 \\ 6 & 0 \end{pmatrix}$ **c.** $\begin{pmatrix} 2 & -2 & -2 \\ 2 & 2 & -2 \\ 2 & 2 & 2 \end{pmatrix}$ **d.** $\begin{pmatrix} 0 & 1 & -1 \\ -1 & 0 & 2 \\ 1 & -2 & 0 \end{pmatrix}$

19. Sean A y B matrices antisimétricas de $n \times n$. Muestre que $A + B$ es antisimétrica.

20. Si A es antisimétrica, muestre que toda componente de la diagonal principal de A es cero.

21. Si A y B son matrices antisimétricas de $n \times n$, muestre que $(AB)^t = BA$, de modo que, AB es simétrica si y sólo si A y B conmutan.

22. Sea $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ una matriz con componentes no negativas y con las propiedades (i) $a_{11}^2 + a_{21}^2 = 1$ y $a_{12}^2 + a_{22}^2 = 1$ y (ii) $\begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix} \cdot \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix} = 0$. Muestre que A es invertible y que $A^{-1} = A^t$.

1.10 Matrices elementales e inversas de matrices

Sea A una matriz de $m \times n$. Entonces, como pronto se verá, podemos efectuar operaciones elementales en renglones o filas sobre A multiplicando A por la izquierda por una matriz apropiada. Las operaciones elementales son

i. Multiplicar el i-ésimo renglón por un número c distinto de cero. $M_i(c)$
ii. Sumar un múltiplo del i-ésimo renglón al j-ésimo renglón. $A_{ij}(c)$
iii. Permutar (intercambiar) los i-ésimo y j-ésimo renglones. P_{ij}

Definición 1 **Matriz elemental.** Una matriz de $n \times n$ (cuadrada) E se denomina *matriz elemental* si puede ser obtenida a partir de la matriz identidad de $n \times n$, I_n, por *una sola* operación elemental de renglones.

Notación. Denotamos una matriz elemental por E o por M_i, A_{ij}, o bien P_{ij}, dependiendo de cómo se obtuvo la matriz a partir de I.

Ejemplo 1 Se obtienen las tres matrices elementales.

(a) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{M_2(5)} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 1 \end{pmatrix} = M_2$ Se tiene multiplicando por 5 el segundo renglón de I.

(b) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{A_{1,3}(-3)} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{pmatrix} = A_{13}$ Se tiene multiplicando por -3 el primer renglón de I y sumándola al tercero.

(c) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{P_{23}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} = P_{23}$ Se tiene permutando los renglones segundo y tercero de I.

La demostración del siguiente teorema se deja como ejercicio (vea los Problemas 58 a 60).

Teorema 1 Para efectuar una operación elemental en renglones sobre la matriz A de $m \times n$, multiplíquese la matriz A, por la izquierda, por la correspondiente matriz elemental.

Ejemplo 2 Sea $A = \begin{pmatrix} 1 & 3 & 2 & 1 \\ 4 & 2 & 3 & -5 \\ 3 & 1 & -2 & 4 \end{pmatrix}$. Efectuar las siguientes operaciones elementales por filas o renglones sobre A, multiplicando A por la izquierda por la matriz elemental apropiada.

(a) Multiplicar el segundo renglón por 5.
(b) Multiplicar el primer renglón por −3 y añadirlo al tercer renglón.
(c) Permutar los renglones segundo y tercero.

Solución Como A es una matriz de 3×4, cada matriz elemental E debe ser de 3×3 puesto que E tiene que ser cuadrada y E está multiplicando por la izquierda a A. Se usan los resultados del Ejemplo 1.

$$\text{(a) } M_2A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 3 & 2 & 1 \\ 4 & 2 & 3 & -5 \\ 3 & 1 & -2 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 3 & 2 & 1 \\ 20 & 10 & 15 & -25 \\ 3 & 1 & -2 & 4 \end{pmatrix}$$

$$\text{(b) } A_{13}A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 3 & 2 & 1 \\ 4 & 2 & 3 & -5 \\ 3 & 1 & -2 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 3 & 2 & 1 \\ 4 & 2 & 3 & -5 \\ 0 & -8 & -8 & 1 \end{pmatrix}$$

$$\text{(c) } P_{23}A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 3 & 2 & 1 \\ 4 & 2 & 3 & -5 \\ 3 & 1 & -2 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 3 & 2 & 1 \\ 3 & 1 & -2 & 4 \\ 4 & 2 & 3 & -5 \end{pmatrix}$$

Considérense los tres productos siguientes:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & c & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/c & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{1}$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ c & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -c & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{2}$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{3}$$

Las Ecuaciones (1), (2) y (3) indican que cada matriz es invertible y que su inversa es del mismo tipo (Tabla 1.4). Esto es consecuencia del Teorema 1. Evidentemente, si las operaciones $A_{ij}(c)$ seguidas de $A_{ij}(-c)$ se efectúan sobre A, la matriz A no se altera. También, $M_i(c)$ seguida de $M_i(1/c)$ y permutando dos veces las mismas filas no alteran la matriz A.

Se tiene que

$$[M_i(c)]^{-1} = M_i\left(\frac{1}{c}\right) \tag{4}$$

$$[A_{ij}(c)]^{-1} = A_{ij}(-c) \tag{5}$$

$$(P_{ij})^{-1} = P_{ij} \tag{6}$$

Tipo de matriz elemental E	Efecto de multiplicar A por la izquierda por E	Representación simbólica de la operación elemental por renglones	Cuando E^{-1} multiplica por la izquierda, tiene el efecto:	Representación simbólica de la operación inversa
Multiplicación	Multiplicación del i-ésimo renglón de A por $c \neq 0$	$M_i(c)$	Multiplicación del i-ésimo renglón de A por $\frac{1}{c}$	$M_i\left(\frac{1}{c}\right)$
Adición	Multiplicación del i-ésimo renglón de A por c y se suma al j-ésimo renglón	$A_{i,j}(c)$	Multiplicación del i-ésimo renglón de A por $-c$ y se suma al j-ésimo renglón	$A_{i,j}(-c)$
Permutación	Permutación de los i-ésimo y j-ésimo renglones de A	P_{ij}	Permutación de los i-ésimo y j-ésimo renglones de A	P_{ij}

La Ecuación (6) indica que

> Toda matriz de permutación es su propia inversa.

En resumen:

Teorema 2 Toda matriz elemental es invertible. La inversa de una matriz elemental es una matriz del mismo tipo.

Teorema 3 Una matriz cuadrada es invertible si y sólo si es el producto de matrices elementales.

Demostración Sea $A = E_1E_2 \cdots E_m$, en donde cada E_i es elemental. Por el Teorema 2, cada E_i es invertible. Más aún, por el Teorema 1.8.3, A es invertible* y

$$A^{-1} = E_m^{-1}E_{m-1}^{-1} \cdots E_2^{-1}E_1^{-1}$$

Recíprocamente, supóngase que A es invertible. Conforme al Teorema 1.8.6 (el Teorema Resumen), A es equivalente por renglones a la matriz identidad. Esto significa que A puede ser reducida a I por un número finito, m, de operaciones elementales por renglones. Por el Teorema 1, cada operación tal puede

* Aquí hemos usado la generalización del Teorema 1.8.3 a más de dos matrices. Vea, por ejemplo, el Problema 1.8.16.

ser efectuada multiplicando A por la izquierda por una matriz elemental. Esto significa que existen matrices elementales $E_1, E_2, \cdots, E_m$ tales que

$$E_m E_{m-1} \cdots E_2 E_1 A = I$$

Así, del Teorema 1.8.8, se desprende que

$$E_m E_{m-1} \cdots E_2 E_1 = A^{-1}$$

y, como cada E_i es invertible, del Teorema 2, resulta que

$$A = (A^{-1})^{-1} = (E_m E_{m-1} \cdots E_2 E_1)^{-1} = E_1^{-1} E_2^{-1} \cdots E_{m-1}^{-1} E_m^{-1} \tag{7}$$

Dado que la inversa de una matriz elemental es también una matriz elemental, se ha logrado expresar a A como el producto de matrices elementales. La demostración queda completa. ■

Ejemplo 4 Demostrar que la matriz $A = \begin{pmatrix} 2 & 4 & 6 \\ 4 & 5 & 6 \\ 3 & 1 & -2 \end{pmatrix}$ es invertible y expresarla como producto de matrices elementales.

Solución Se ha encontrado esta matriz anteriormente, primero en el Ejemplo 1.6.1. Para resolver el problema, se reduce A a I y se sigue el rastro de las operaciones elementales sobre renglones. En el Ejemplo 1.8.6 se redujo A a I por las operaciones siguientes:

① $M_1(\frac{1}{2})$, ② $A_{1,2}(-4)$, ③ $A_{1,3}(-3)$, ④ $M_2(-\frac{1}{3})$,

⑤ $A_{2,1}(-2)$, ⑥ $A_{2,3}(5)$, ⑦ $M_3(-1)$,

⑧ $A_{3,1}(1)$, ⑨ $A_{3,2}(-2)$.

A^{-1} se obtuvo comenzando con I y aplicando estas nueve operaciones elementales. Así pues, A^{-1} es el producto de nueve matrices elementales:

$$A^{-1} = \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}}_{A_{3,2}(-2)} \underbrace{\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{A_{3,1}(1)} \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}}_{M_3(-1)} \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 5 & 1 \end{pmatrix}}_{A_{2,3}(5)} \underbrace{\begin{pmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{A_{2,1}(-2)}$$

$$\times \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{3} & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{M_2(-\frac{1}{3})} \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{pmatrix}}_{A_{1,3}(-3)} \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ -4 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{A_{1,2}(-4)} \underbrace{\begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{M_1(\frac{1}{2})}$$

Entonces $A = (A^{-1})^{-1}$ = producto de las inversas de las nueve matrices en el orden contrario:

$$\begin{pmatrix} 2 & 4 & 6 \\ 4 & 5 & 6 \\ 3 & 1 & -2 \end{pmatrix} = \underbrace{\begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{M_1(2)} \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{A_{1,2}(4)} \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 0 & 1 \end{pmatrix}}_{A_{1,3}(3)} \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 0 \end{pmatrix}}_{M_2(-3)} \underbrace{\begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{A_{2,1}(2)}$$

$$\times \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -5 & 1 \end{pmatrix}}_{A_{2,3}(-5)} \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}}_{M_3(-1)} \underbrace{\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{A_{3,1}(-1)} \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}}_{A_{3,2}(2)}$$

Podemos usar el Teorema 3 para extender el Teorema Resumen.

Teorema 4 ***Teorema Resumen — Versión 3.*** Sea A una matriz de $n \times n$. Entonces cada uno de los seis enunciados siguientes implica a los otros cinco (esto es, si uno de ellos es verdadero, todos son verdaderos):

i. A es invertible.
ii. La única solución al sistema homogéneo $A\mathbf{x} = \mathbf{0}$ es la solución trivial ($\mathbf{x} = \mathbf{0}$).
iii. El sistema $A\mathbf{x} = \mathbf{b}$ posee una solución única para cada n-vector $\mathbf{b}$.
iv. A es equivalente por renglones a la matriz identidad I_n de $n \times n$.
v. A puede expresarse como producto de matrices elementales.
vi. $\det A \neq 0$ (hasta ahora, $\det A$ sólo se ha definido si A es una matriz de 2×2).

Concluimos esta sección demostrando la parte (*iv*) del Teorema 1.9.1.

Teorema 5 Sea E una matriz elemental. Entonces E^t es invertible y

$$(E^t)^{-1} = (E^{-1})^t \tag{8}$$

Demostración ***Caso 1:*** E es una $M_i(c)$. Entonces $M_i(c)$ es simétrica ($E^t = E$), al igual que lo es E^{-1}, de modo que (8) se verifica.

Caso 2: E se representa por $A_{ij}(c)$. Supóngase que E es 3×3 y que $i = 1$, $j = 2$. Entonces

$$E = \begin{pmatrix} 1 & 0 & 0 \\ c & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad E^t = \begin{pmatrix} 1 & c & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad E^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -c & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Es fácil verificar que

$$(E^t)^{-1} = \begin{pmatrix} 1 & -c & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = (E^{-1})^t$$

En el caso más general, no es difícil comprobar que si E es una matriz con ele-

mentos 1 en la diagonal, una c en la ij-ésima posición, y 0 en todos los demás lugares, entonces E^{-1} tiene el valor $-c$ en la ij-ésima posición, E^t tiene el valor c en la ij-ésima posición, y ambas, $(E^{-1})^t$ y $(E^t)^{-1}$ tienen $-c$ en la ij-ésima posición, elementos 1 en la diagonal principal y ceros en todas las demás posiciones. En suma,

$$[(A_{ij}(c))^{-1}]^t = A_{ji}(-c) = [(A_{ij}(c))^t]^{-1}$$

Caso 3: E es una matriz de permutación. Entonces $E = E^{-1}$, así que $E^t = (E^{-1})^t$. Así, E^t representa el proceso de intercambiar las i-ésima y j-ésima columnas de la matriz A. Esto significa que E^t es su propia inversa:

$$(E^t)^{-1} = E^t = (E^{-1})^t \quad ■$$

Teorema 6 Si A es una matriz invertible $n \times n$, entonces A^t es invertible y

$$(A^t)^{-1} = (A^{-1})^t \qquad (9)$$

Demostración Por el Teorema 3, A es el producto de matrices elementales:

$$A = E_1E_2 \cdots E_m \qquad (10)$$

Aplicando el Teorema 1.8.3,

$$A^{-1} = E_m^{-1}E_{m-1}^{-1} \cdots E_1^{-1} \qquad (11)$$

y por la Ecuación (1.9.3),

$$A^t = E_m^tE_{m-1}^t \cdots E_1^t \qquad (12)$$

De (11) y de la Ecuación (1.9.3),

$$(A^{-1})^t = (E_1^{-1})^t(E_2^{-1})^t \cdots (E_m^{-1})^t \qquad (13)$$

De (12) y del Teorema 1.8.3,

$$(A^t)^{-1} = (E_1^t)^{-1}(E_2^t)^{-1} \cdots (E_m^t)^{-1} \qquad (14)$$

Por el Teorema 5, los productos en (13) y (14) son iguales. La demostración queda completa. ■

Ejemplo 5 Sea $A = \begin{pmatrix} 2 & 4 & 6 \\ 4 & 5 & 6 \\ 3 & 1 & -2 \end{pmatrix}$. Entonces

$$A^t = \begin{pmatrix} 2 & 4 & 3 \\ 4 & 5 & 1 \\ 6 & 6 & -2 \end{pmatrix} \quad \text{y} \quad A^{-1} \overset{\text{Ejemplo 1.8.6,}}{=} \frac{1}{6}\begin{pmatrix} -16 & 14 & -6 \\ 26 & -22 & 12 \\ -11 & 10 & -6 \end{pmatrix}$$

Se deja al lector verificar que

$$(A^t)^{-1} = \frac{1}{6}\begin{pmatrix} -16 & 26 & -11 \\ 14 & -22 & 10 \\ -6 & 12 & -6 \end{pmatrix} = (A^{-1})^t$$

Hay un resultado más que será de utilidad en la Sección 2.3.

Teorema 7 Sea A una matriz de $n \times n$. Entonces A puede expresarse como un producto de matrices elementales y de una matriz T triangular superior.

Demostración Al aplicar eliminación gaussiana al sistema $A\mathbf{x} = \mathbf{b}$, obtenemos una matriz triangular superior. Para captar esto, se observa que la eliminación gaussiana termina cuando la matriz queda reducida a la forma escalón por renglones —y esta forma (para una matriz de $n \times n$) es triangular superior. Denotamos la forma escalón de A por T. Entonces, A se reduce a T por una sucesión de operaciones elementales por filas, cada una de las cuales se puede obtener multiplicando por una matriz elemental. Así pues,

$$T = E_mE_{m-1} \cdots E_2E_1A$$

y

$$A = E_1^{-1}E_2^{-1} \cdots E_{m-1}^{-1}E_m^{-1}T$$

Como la inversa de una matriz elemental es una matriz elemental, se ha escrito A como el producto de matrices elementales y T. ■

Ejemplo 6 Expresar la matriz

$$A = \begin{pmatrix} 3 & 6 & 9 \\ 2 & 5 & 1 \\ 1 & 1 & 8 \end{pmatrix}$$

como producto de matrices elementales y una matriz triangular superior.

Solución Se reduce por renglones A a su forma escalonada:

$$\begin{pmatrix} 3 & 6 & 9 \\ 2 & 5 & 1 \\ 1 & 1 & 8 \end{pmatrix} \xrightarrow{M_1(\frac{1}{3})} \begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 1 \\ 1 & 1 & 8 \end{pmatrix} \xrightarrow[A_{1,3}(-1)]{A_{1,2}(-2)}$$

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & -5 \\ 0 & -1 & 5 \end{pmatrix} \xrightarrow{A_{2,3}(1)} \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & -5 \\ 0 & 0 & 0 \end{pmatrix} = T$$

Entonces, procediendo hacia atrás resulta que

$$T = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & -5 \\ 0 & 0 & 0 \end{pmatrix} = \underset{A_{2,3}(1)}{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}} \underset{A_{1,3}(-1)}{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}}$$

$$\times \underset{A_{1,2}(-2)}{\begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}} \underset{M_1(\frac{1}{3})}{\begin{pmatrix} \frac{1}{3} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}} \underset{A}{\begin{pmatrix} 3 & 6 & 9 \\ 2 & 5 & 1 \\ 1 & 1 & 8 \end{pmatrix}}$$

y tomando inversas de las cuatro matrices elementales,

$$A = \begin{pmatrix} 3 & 6 & 9 \\ 2 & 5 & 1 \\ 1 & 1 & 8 \end{pmatrix} = \underset{M_3(3)}{\begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}} \underset{A_{1,2}(2)}{\begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}$$

$$\times \underset{A_{1,3}(1)}{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}} \underset{A_{2,3}(-1)}{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}} \underset{T}{\begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & -5 \\ 0 & 0 & 0 \end{pmatrix}}$$

Problemas 1.10

En los Problemas del 1 *al* 12 *determine qué matrices son elementales.*

1. $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ **2.** $\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ **3.** $\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$

4. $\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$ **5.** $\begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix}$ **6.** $\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

7. $\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$ **8.** $\begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 0 & 1 \end{pmatrix}$ **9.** $\begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

10. $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}$ **11.** $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}$ **12.** $\begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$

En los Problemas del 13 *al* 20 *halle una matriz elemental de* 3 × 3 *que efectúe la operación por renglones indicada sobre una matriz A de* 3 × 5 *al multiplicarse por la izquierda.*

13. $M_2(4)$ **14.** $A_{1,2}(2)$ **15.** $A_{2,1}(-3)$ **16.** $A_{3,1}(4)$
17. P_{13} **18.** P_{23} **19.** $A_{3,2}(1)$ **20.** $M_3(-1)$

En los Problemas del 21 *al* 30, *obtenga la matriz elemental E tal que EA = B.*

21. $A = \begin{pmatrix} 2 & 3 \\ -1 & 4 \end{pmatrix}, B = \begin{pmatrix} 2 & 3 \\ 2 & -8 \end{pmatrix}$ **22.** $A = \begin{pmatrix} 2 & 3 \\ -1 & 4 \end{pmatrix}, B = \begin{pmatrix} 2 & 3 \\ -5 & -2 \end{pmatrix}$

23. $A = \begin{pmatrix} 2 & 3 \\ -1 & 4 \end{pmatrix}, B = \begin{pmatrix} 0 & 11 \\ -1 & 4 \end{pmatrix}$ **24.** $A = \begin{pmatrix} 2 & 3 \\ -1 & 4 \end{pmatrix}, B = \begin{pmatrix} -1 & 4 \\ 2 & 3 \end{pmatrix}$

25. $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}, B = \begin{pmatrix} 5 & 6 \\ 3 & 4 \\ 1 & 2 \end{pmatrix}$ **26.** $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}, B = \begin{pmatrix} 1 & 2 \\ 0 & -2 \\ 5 & 6 \end{pmatrix}$

27. $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}, B = \begin{pmatrix} -1 & -2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}$ **28.** $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}, B = \begin{pmatrix} -5 & -6 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}$

29. $A = \begin{pmatrix} 1 & 2 & 5 & 2 \\ 0 & -1 & 3 & 4 \\ 5 & 0 & -2 & 7 \end{pmatrix}, B = \begin{pmatrix} 1 & 2 & 5 & 2 \\ 0 & -1 & 3 & 4 \\ 0 & -10 & -27 & -3 \end{pmatrix}$

30. $A = \begin{pmatrix} 1 & 2 & 5 & 2 \\ 0 & -1 & 3 & 4 \\ 5 & 0 & -2 & 7 \end{pmatrix}, B = \begin{pmatrix} 1 & 0 & 11 & 10 \\ 0 & -1 & 3 & 4 \\ 5 & 0 & -2 & 7 \end{pmatrix}$

En los Problemas del 31 *al* 40 *halle la inversa de la matriz elemental dada.*

31. $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ **32.** $\begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}$ **33.** $\begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix}$

34. $\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ **35.** $\begin{pmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ **36.** $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix}$

37. $\begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}$ **38.** $\begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ **39.** $\begin{pmatrix} 1 & 0 & 0 & 5 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$

40. $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -3 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$

En los Problemas del 41 *al* 48 *demuestre que cada matriz es invertible y exprésela como producto de matrices elementales.*

41. $\begin{pmatrix} 2 & 1 \\ 3 & 2 \end{pmatrix}$ **42.** $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ **43.** $\begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 3 \\ 5 & 5 & 1 \end{pmatrix}$

44. $\begin{pmatrix} 3 & 2 & 1 \\ 0 & 2 & 2 \\ 0 & 0 & -1 \end{pmatrix}$ **45.** $\begin{pmatrix} 0 & -1 & 0 \\ 0 & 1 & -1 \\ 1 & 0 & 1 \end{pmatrix}$ **46.** $\begin{pmatrix} 2 & 0 & 4 \\ 0 & 1 & 1 \\ 3 & -1 & 1 \end{pmatrix}$

47. $\begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & -4 & 0 \\ 0 & 0 & 0 & 5 \end{pmatrix}$ **48.** $\begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}$

49. Sea $A = \begin{pmatrix} a & b \\ 0 & c \end{pmatrix}$, en donde $ac \neq 0$. Escriba A como producto de tres matrices elementales y concluya que A es invertible.

50. Sea $A = \begin{pmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & f \end{pmatrix}$, en donde $adf \neq 0$. Exprese A como producto de seis matrices elementales y concluya que A es invertible.

★ 51. Sea A una matriz $n \times n$ triangular superior (vea el Problema 1.8.29). Demuestre que si no es cero cada componente en la diagonal, entonces A es invertible. [*Sugerencia:* Vea los Problemas 49 y 50.]

★ 52. Demuestre que si A es una matriz triangular superior de $n \times n$, con componentes diagonales no nulas, entonces A^{-1} también es triangular superior.

★ 53. Aplique el Teorema 6 y el resultado del Problema 52 para demostrar que si A es una matriz de $n \times n$ triangular inferior con componentes diagonales distintas de cero, entonces A es invertible y A^{-1} es triangular inferior.

En los Problemas del 54 *al* 57 *evalúe* $(A^t)^{-1}$ *y* $(A^{-1})^t$, *y demuestre que son iguales.*

54. $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$

55. $A = \begin{pmatrix} 2 & 1 \\ 3 & 2 \end{pmatrix}$

56. $A = \begin{pmatrix} 3 & 2 & 1 \\ 0 & 2 & 2 \\ 0 & 0 & -1 \end{pmatrix}$

57. $A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 3 \\ 5 & 5 & 1 \end{pmatrix}$

58. Demuestre que si P_{ij} es la matriz de $n \times n$ obtenida permutando los i-ésimo y j-ésimo renglones de I_n, entonces $P_{ij}\,A$ es la matriz que se obtiene de A al permutar sus i-ésimo y j-ésimo renglones.

59. Sea A_{ij} la matriz con c en la ij-ésima posición, elementos 1 en la diagonal, y ceros en el resto de las posiciones. Demuestre que $A_{ij}\,A$ es la matriz que se obtiene de A al multiplicar por c el i-ésimo renglón de A y sumarla al j-ésimo renglón.

60. Sea M_i la matriz con c en la posición ii, elementos 1 en las otras posiciones diagonales y ceros en el resto de las posiciones. Demuestre que $M_i\,A$ es la matriz que se obtiene de A multiplicando por c su i-ésimo renglón.

En los Problemas del 61 *al* 66 *exprese cada matriz cuadrada como producto de matrices elementales y de una matriz triangular superior.*

61. $A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$

62. $A = \begin{pmatrix} 2 & -3 \\ -4 & 6 \end{pmatrix}$

63. $A = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$

64. $A = \begin{pmatrix} 1 & -1 & 2 \\ 2 & 1 & 4 \\ 4 & -1 & 8 \end{pmatrix}$

65. $A = \begin{pmatrix} 1 & -3 & 3 \\ 0 & -3 & 1 \\ 1 & 0 & 2 \end{pmatrix}$

66. $A = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 3 & 0 \\ -1 & 4 & 0 \end{pmatrix}$

1.11 Una perspectiva diferente: inversas unilaterales de matrices

En la Sección 1.8 definimos la inversa de una matriz cuadrada. Se demostró que el sistema $A\mathbf{x} = \mathbf{b}$ tiene solución única si y sólo si A es invertible. ¿Qué sucede empero si A no es cuadrada? ¿Posee soluciones el sistema $A\mathbf{x} = \mathbf{b}$? ¿Tiene soluciones únicas? Daremos respuestas parciales a estas preguntas en esta sección.

Definición 1 ***Inversas izquierda y derecha.*** Sea A una matriz de $m \times n$. Entonces

i. A posee *inversa derecha* si existe una matriz R de $n \times m$ tal que

$$AR = I_m \tag{1}$$

ii. A posee *inversa izquierda* si existe una matriz L de $n \times m$ tal que

$$LA = I_n \tag{2}$$

Observación 1. AR tiene dimensiones $(m \times n) \times (n \times m) = m \times m$.
LA tiene dimensiones $(n \times m) \times (m \times n) = n \times n$.
Así que las matrices identidad en (1) y (2) son desiguales si $m \neq n$.

Observación 2. Si A es cuadrada, entonces, por el Teorema 1.8.8, A es invertible si y sólo si A tiene inversas izquierda y derecha. En este caso, $L = R = A^{-1}$. Por esta razón, sólo se consideran en esta sección matrices no cuadradas; esto es, $m \neq n$.

Antes de calcular las inversas derechas o izquierdas, demostraremos varios teoremas. Consideremos el sistema

$$A\mathbf{x} = \mathbf{b} \tag{3}$$

en donde A es $m \times n$; $\mathbf{x}$ y $\mathbf{b}$ son vectores de dimensiones n y m, respectivamente.

Teorema 1 ***Teorema de existencia.*** La matriz A de $m \times n$ tiene inversa derecha si y sólo si el sistema (3) posee al menos una solución para cada vector $\mathbf{b}$ m-dimensional.

Demostración Supongamos primero que A tiene inversa derecha R. Defínase $\mathbf{x}^* = R\mathbf{b}$. Nótese que $\mathbf{x}^*$ es $n \times 1$ puesto que R es $n \times m$ y $\mathbf{b}$ es $m \times 1$. Entonces

$$A\mathbf{x}^* = A(R\mathbf{b}) = (AR)\mathbf{b} = I_m\mathbf{b} = \mathbf{b}$$

Así pues, $\mathbf{x}^*$ es solución de (3).

Recíprocamente, si (3) posee una solución para todo m-vector $\mathbf{b}$, sean

$$\mathbf{b}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \mathbf{b}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \mathbf{b}_j = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \mathbf{b}_m = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

j-ésima posición

Cada $\mathbf{b}_j$ es un m-vector o vector-m.
Sea

$$\mathbf{r}_j = \begin{pmatrix} r_{1j} \\ r_{2j} \\ \vdots \\ r_{nj} \end{pmatrix}$$

una solución de $A\mathbf{x} = \mathbf{b}_j$: y finalmente,

$$R = \begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1j} & \cdots & r_{1m} \\ r_{21} & r_{22} & \cdots & r_{2j} & \cdots & r_{2m} \\ \vdots & \vdots & & \vdots & & \vdots \\ r_{n1} & r_{n2} & \cdots & r_{nj} & \cdots & r_{nm} \end{pmatrix}$$

Demostraremos que R es inversa derecha. Denotemos por $\mathbf{a}_i$ el i-ésimo renglón de A. Entonces, como

$$A\mathbf{r}_j = \mathbf{b}_j$$

tenemos que

$$\mathbf{a}_i \cdot \mathbf{r}_j = i\text{-ésima componente de } \mathbf{b}_j = \begin{cases} 0, \text{ si } i \neq j \\ 1, \text{ si } i = j \end{cases} \tag{4}$$

Pero $\mathbf{a}_i \cdot \mathbf{r}_j$ es la ij-ésima componente de AR. De (4) vemos que $AR = I_m$. ■

Corolario Si R es inversa derecha de A, entonces

$$\mathbf{x} = R\mathbf{b} \tag{5}$$

es una solución de $A\mathbf{x} = \mathbf{b}$.

Teorema 2 ***Teorema de unicidad.*** Si la matriz A de $m \times n$ posee inversa izquierda, entonces el sistema (3) tiene, cuando más una sola solución para todo m-vector $\mathbf{b}$.

Demostración Sean $\mathbf{x}$, $\mathbf{y}$ soluciones de (3). Entonces

$$A\mathbf{x} = \mathbf{b} \text{ y también } A\mathbf{y} = \mathbf{b}$$

así que

$$A(\mathbf{x} - \mathbf{y}) = A\mathbf{x} - A\mathbf{y} = \mathbf{b} - \mathbf{b} = \mathbf{0}$$

Si ahora L es inversa izquierda de A,

$$(\mathbf{x} - \mathbf{y}) = I_n(\mathbf{x} - \mathbf{y}) = LA(\mathbf{x} - \mathbf{y}) = L\mathbf{0} = \mathbf{0}$$

de modo que $\mathbf{x} = \mathbf{y}$. ■

Teorema 3 Sea A una matriz de $m \times n$. Entonces

i. A posee inversa izquierda si y sólo si A^t tiene una inversa derecha.
ii. Si $n > m$, A no tiene inversa izquierda.
iii. Si $m > n$, A no tiene inversa derecha.

Demostración **i.** Supongamos que A posee inversa izquierda. Entonces existe una matriz L de $n \times m$, tal que $LA = I_n$. Transponiendo, queda

$$I_n \overset{I_n \text{ es simétrica}}{=} I_n^t = (LA)^t \overset{\text{Teorema 1.9.1(ii)}}{=} A^tL^t$$

lo cual muestra que L^t es inversa derecha de A^t. Similarmente, si A tiene inversa derecha R de $n \times m$, entonces, $AR = I_m$ y

$$I_m = I_m^t = (AR)^t = R^tA^t$$

por lo que R^t es inversa izquierda de A^t.

ii. Por el Teorema 1.7.1, el sistema $A\mathbf{x} = \mathbf{0}$ tiene un número infinito de soluciones, ya que $n > m$. De manera que A no puede poseer inversa izquierda, pues si así fuera, el sistema $A\mathbf{x} = \mathbf{0}$ sólo tendría una solución, según el Teorema 2.

iii. Si A tuviera inversa derecha, entonces A^t tendría inversa izquierda, por la parte (i) del teorema. Pero A^t es una matriz de $n \times m$ con $m > n$; se tiene entonces que A^t no posee inversa izquierda, con base en la parte (ii). En consecuencia, A no posee inversa derecha. ■

Corolario. Sea A una matriz de $m \times n$ con $m \neq n$. Entonces el sistema $A\mathbf{x} = \mathbf{b}$ no tiene solución única para todo n-vector $\mathbf{b}$.

Demostración ***Caso 1:*** $m > n$. Por el Teorema 3(iii), A no posee inversa derecha, y por lo mismo, por el Teorema 1, $A\mathbf{x} = \mathbf{b}$ no tiene solución para todo m-vector $\mathbf{b}$.

Caso 2: $n > m$. Como en la demostración del Teorema 3(ii), $A\mathbf{x} = \mathbf{0}$ tiene un número infinito de soluciones. ■

Ejemplo 1 Encontrar, si existe, una inversa derecha de $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$

Solución Buscamos una matriz $R = \begin{pmatrix} u & v \\ w & x \\ y & z \end{pmatrix}$ tal que $AR = I_2$. Esto es,

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}\begin{pmatrix} u & v \\ w & x \\ y & z \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{6}$$

Escribimos el sistema (7) como una matriz aumentada, y se reduce por filas o renglones:

$$\begin{array}{cccccc} u & v & w & x & y & z \end{array}$$
$$\left(\begin{array}{cccccc|c} 1 & 0 & 2 & 0 & 3 & 0 & 1 \\ 0 & 1 & 0 & 2 & 0 & 3 & 0 \\ 4 & 0 & 5 & 0 & 6 & 0 & 0 \\ 0 & 4 & 0 & 5 & 0 & 6 & 1 \end{array}\right) \xrightarrow{A_{1,3}(-4)} \left(\begin{array}{cccccc|c} 1 & 0 & 2 & 0 & 3 & 0 & 1 \\ 0 & 1 & 0 & 2 & 0 & 3 & 0 \\ 0 & 0 & -3 & 0 & -6 & 0 & -4 \\ 0 & 4 & 0 & 5 & 0 & 6 & 1 \end{array}\right)$$

$$\xrightarrow{A_{2,4}(-4)}\left(\begin{array}{cccccc|c} 1 & 0 & 2 & 0 & 3 & 0 & 1 \\ 0 & 1 & 0 & 2 & 0 & 3 & 0 \\ 0 & 0 & -3 & 0 & -6 & 0 & -4 \\ 0 & 0 & 0 & -3 & 0 & -6 & 1 \end{array}\right)$$

$$\xrightarrow[M_4(-\frac{1}{3})]{M_3(-\frac{1}{3})}\left(\begin{array}{cccccc|c} 1 & 0 & 2 & 0 & 3 & 0 & 1 \\ 0 & 1 & 0 & 2 & 0 & 3 & 0 \\ 0 & 0 & 1 & 0 & 2 & 0 & \frac{4}{3} \\ 0 & 0 & 0 & 1 & 0 & 2 & -\frac{1}{3} \end{array}\right) \xrightarrow[A_{4,2}(-2)]{A_{3,1}(-2)} \begin{array}{c} \begin{array}{cccccc} u & v & w & x & y & z \end{array} \\ \left(\begin{array}{cccccc|c} 1 & 0 & 0 & 0 & -1 & 0 & -\frac{5}{3} \\ 0 & 1 & 0 & 0 & 0 & -1 & \frac{2}{3} \\ 0 & 0 & 1 & 0 & 2 & 0 & \frac{4}{3} \\ 0 & 0 & 0 & 1 & 0 & 2 & -\frac{1}{3} \end{array}\right) \end{array}$$

Por lo tanto,

$$u = y - \tfrac{5}{3}$$
$$v = z + \tfrac{2}{3}$$
$$w = -2y + \tfrac{4}{3}$$
$$x = -2z - \tfrac{1}{3}$$

Evidentemente, existe una infinidad de soluciones. Eligiendo $y = z = 0$,

$$R_1 = \begin{pmatrix} -\frac{5}{3} & \frac{2}{3} \\ \frac{4}{3} & -\frac{1}{3} \\ 0 & 0 \end{pmatrix}$$

Verificación

$$AR_1 = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}\begin{pmatrix} -\frac{5}{3} & \frac{2}{3} \\ \frac{4}{3} & -\frac{1}{3} \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Eligiendo $y = z = 1$, obtenemos otra inversa derecha:

$$R_2 = \begin{pmatrix} -\frac{2}{3} & \frac{5}{3} \\ -\frac{2}{3} & -\frac{7}{3} \\ 1 & 1 \end{pmatrix}$$

Consideremos el sistema $A\mathbf{x} = \begin{pmatrix} 5 \\ 2 \end{pmatrix}$. De acuerdo con (5), dos soluciones son

$$\mathbf{x}_1 = R_1\mathbf{b} = \frac{1}{3}\begin{pmatrix} -5 & 2 \\ 4 & -1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 5 \\ 2 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} -21 \\ 18 \\ 0 \end{pmatrix} = \begin{pmatrix} -7 \\ 6 \\ 0 \end{pmatrix}$$

y

$$\mathbf{x}_2 = R_2\mathbf{b} = \frac{1}{3}\begin{pmatrix} -2 & 5 \\ -2 & -7 \\ 3 & 3 \end{pmatrix}\begin{pmatrix} 5 \\ 2 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 0 \\ -24 \\ 21 \end{pmatrix} = \begin{pmatrix} 0 \\ -8 \\ 7 \end{pmatrix}$$

Verificación

$$A\mathbf{x}_1 = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}\begin{pmatrix} -7 \\ 6 \\ 0 \end{pmatrix} = \begin{pmatrix} 5 \\ 2 \end{pmatrix} \quad \text{y} \quad A\mathbf{x}_2 = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}\begin{pmatrix} 0 \\ -8 \\ 7 \end{pmatrix} = \begin{pmatrix} 5 \\ 2 \end{pmatrix}$$

Sin hacer más cálculos, concluimos que A no tiene inversa izquierda, pues de lo contrario, se llegaría a una contradicción respecto al resultado del corolario del Teorema 3.

El teorema siguiente generaliza el resultado del Ejemplo 1. Su demostración se deja como ejercicio (vea Problema 16).

Teorema 4 Si la matriz A de $m \times n$ $(m \neq n)$ posee inversa derecha, entonces tiene un número infinito de inversas derechas.

Ejemplo 2 Encontrar una inversa izquierda de

$$A = \begin{pmatrix} -2 & 5 \\ -2 & -7 \\ 3 & 3 \end{pmatrix}$$

Solución Del Ejemplo 1, vemos que

$$L = \frac{1}{3}\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$$

Verificación

$$LA = \frac{1}{3}\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}\begin{pmatrix} -2 & 5 \\ -2 & -7 \\ 3 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Ejemplo 3 Consideremos el sistema $A\mathbf{x} = \mathbf{b}$, en donde $A = \begin{pmatrix} 1 & 2 \\ 2 & 5 \\ 3 & 4 \end{pmatrix}$. A posee la inversa izquierda $L = \begin{pmatrix} -2 & 0 & 1 \\ 5 & -1 & -1 \end{pmatrix}$. Así que A no tiene inversa derecha y el sistema carece de solución para ningún vector tridimensional $\mathbf{b}$. Para ver esto más claramente, consideremos $\mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$ y reduzcamos por renglones:

$$\left(\begin{array}{cc|c} 1 & 2 & b_1 \\ 2 & 5 & b_2 \\ 3 & 4 & b_3 \end{array}\right) \xrightarrow[A_{1,3}(-3)]{A_{1,2}(-2)} \left(\begin{array}{cc|c} 1 & 2 & b_1 \\ 0 & 1 & -2b_1 + b_2 \\ 0 & -2 & -3b_1 + b_3 \end{array}\right)$$

$$\xrightarrow[A_{2,3}(2)]{A_{2,1}(-2)} \left(\begin{array}{cc|c} 1 & 0 & 5b_1 - 2b_2 \\ 0 & 1 & -2b_1 + b_2 \\ 0 & 0 & -7b_1 + 2b_2 + b_3 \end{array}\right)$$

Evidentemente el sistema es inconsistente, a menos que $-7b_1 + 2b_2 + b_3 = 0$. Por ejemplo, si $\mathbf{b} = \begin{pmatrix} 3 \\ 8 \\ 5 \end{pmatrix}$ entonces $-7b_1 + 2b_2 + b_3 = 0$ y el sistema posee la solución única $\begin{pmatrix} 5b_1 - 2b_2 \\ -2b_1 + b_2 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$. Por otro lado, si $\mathbf{b} = \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix}$, entonces $-7b_1 + 2b_2 + b_3 = -4 \neq 0$ y el sistema no tiene solución.

Podemos apreciar la existencia de matrices unilaterales de una matriz desde otro punto de vista. Si A es una matriz de $m \times n$, es posible contemplar A como representante de una *función* cuyo dominio es el conjunto de n-vectores y cuyo ámbito (o contradominio) es el conjunto de los m-vectores.

La función $\mathbf{b} = A\mathbf{x}$ es *biyectiva* si, siempre que $A\mathbf{x} = A\mathbf{y}$, es posible concluir que $\mathbf{x} = \mathbf{y}$; esto es, si toda $\mathbf{b}$ en el ámbito es imagen de un solo n-vector.

La función $\mathbf{b} = A\mathbf{x}$ es *suprayectiva* (o *sobre*) respecto del conjunto de m-vectores si todo m-vector $\mathbf{b}$ en el ámbito es imagen de un n-vector del dominio. Esto es, si existe x tal que $\mathbf{b} = A\mathbf{x}$.

Podemos enunciar los Teoremas 1 y 2 del modo siguiente:

Teorema 1' Sea A una matriz de $m \times n$. La función $\mathbf{b} = A\mathbf{x}$ es suprayectiva respecto del conjunto de m-vectores si y sólo si A tiene inversa derecha.

Teorema 2' Sea A una matriz de $m \times n$. Si A posee inversa izquierda, entonces la función $\mathbf{b} = A\mathbf{x}$ es biyectiva.

Se expresará mucho más acerca de funciones biyectivas (o biunívocas) y suprayectivas (o sobre) cuando se analicen transformaciones lineales en el Capítulo 5 (vea, en particular, la Sección 5.4).

Problemas 1.11

En los Problemas del 1 *al* 10 *determine cuáles matrices poseen inversas derechas o izquierdas. Si existen, calcular una.*

1. $\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$ **2.** $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$ **3.** $\begin{pmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \end{pmatrix}$ **4.** $\begin{pmatrix} 2 & -1 & 5 \\ 3 & 4 & 1 \end{pmatrix}$

5. $\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}$ **6.** $\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$ **7.** $\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}$ **8.** $\begin{pmatrix} 3 & -2 \\ 4 & 1 \\ 7 & 2 \end{pmatrix}$

9. $(1 \quad 2 \quad 3)$ **10.** $\begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix}$

11. Sea $A = \begin{pmatrix} 2 & 3 \\ -1 & 4 \\ 2 & 5 \end{pmatrix}$

a. Demuestre que A no tiene inversa derecha.

b. Halle una condición que asegure una solución para el sistema $A\mathbf{x} = \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$.

12. Contestar las preguntas del Problema 11 para la matriz

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & -1 \\ 0 & 2 & 4 \\ 3 & 6 & 2 \end{pmatrix}$$

13. **a.** Obtenga dos inversas derechas de $A = \begin{pmatrix} 1 & 3 & 0 \\ 4 & 0 & 2 \end{pmatrix}$.

b. Use los resultados de (a) para determinar dos soluciones al sistema $A\mathbf{x} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

14. **a.** Obtenga dos inversas derechas de $A = \begin{pmatrix} 1 & -1 & 2 \\ 2 & 4 & 5 \end{pmatrix}$.

b. Use los resultados de (a) para hallar dos soluciones al sistema $A\mathbf{x} = \begin{pmatrix} 2 \\ -3 \end{pmatrix}$.

15. Supóngase que la matriz A de $m \times n$ poseyera inversa derecha R e inversa izquierda L. Demuestre que se tendría que $R = L$.

★ 16. Supóngase que la matriz A de $m \times n$ tiene inversa derecha R con $m \neq n$. Demuestre que A tiene un número infinito de inversas derechas. [*Sugerencia:* Considere el sistema $AR = I$ y demuestre que si este sistema tiene una solución única, entonces lo mismo puede decirse del sistema $A\mathbf{x} = \mathbf{b}$ para todo todo m-vector $\mathbf{b}$, lo cual contradice al Teorema 3.]

17. Considere el sistema $A\mathbf{x} = \mathbf{b}$ donde

$$A = \begin{pmatrix} 2 & -5 \\ 2 & 7 \\ -3 & -3 \end{pmatrix}$$

a. Demuestre que $L = -\frac{1}{3}\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$ es inversa izquierda de A.

b. Sea $\mathbf{b} = \begin{pmatrix} 3 \\ -1 \\ 2 \end{pmatrix}$. Calcule $L\mathbf{b}$.

c. Demuestre que $L\mathbf{b}$ *no* es solución del sistema $A\mathbf{x} = \mathbf{b}$.

d. Multiplique por L ambos lados de $A\mathbf{x} = \mathbf{b}$ por la izquierda. Resulta así

$$A\mathbf{x} = \mathbf{b}$$
$$LA\mathbf{x} = L\mathbf{b}$$

$$I\mathbf{x} = L\mathbf{b}$$

$$\mathbf{x} = L\mathbf{b}$$

Parecería que $\mathbf{x} = L\mathbf{b}$ debiera ser solución de $A\mathbf{x} = \mathbf{b}$. Pero esto contradice lo hallado en (c). De hecho, si $A\mathbf{x} = \mathbf{b}$ tiene una solución para cada $\mathbf{b}$, entonces A posee inversa derecha. Esto contradice el corolario del Teorema 3. ¿Dónde está el error en este razonamiento?

Ejercicios de repaso • Capítulo 1

En los Ejercicios del 1 al 14, obtener todas las soluciones (si existen) de los sistemas dados.

1. $\begin{aligned} 3x_1+6x_2&=9 \\ -2x_1+3x_2&=4 \end{aligned}$

2. $\begin{aligned} 3x_1+6x_2&=9 \\ 2x_1+4x_2&=6 \end{aligned}$

3. $\begin{aligned} 3x_1-6x_2&=9 \\ -2x_1+4x_2&=6 \end{aligned}$

4. $\begin{aligned} x_1+x_2+x_3&=2 \\ 2x_1-x_2+2x_3&=4 \\ -3x_1+2x_2+3x_3&=8 \end{aligned}$

5. $\begin{aligned} x_1+x_2+x_3&=0 \\ 2x_1-x_2+2x_3&=0 \\ -3x_1+2x_2+3x_3&=0 \end{aligned}$

6. $\begin{aligned} x_1+x_2+x_3&=2 \\ 2x_1-x_2+2x_3&=4 \\ -x_1+4x_2+x_3&=2 \end{aligned}$

7. $\begin{aligned} x_1+x_2+x_3&=2 \\ 2x_1-x_2+2x_3&=4 \\ -x_1+4x_2+x_3&=3 \end{aligned}$

8. $\begin{aligned} x_1+x_2+x_3&=0 \\ 2x_1-x_2+2x_3&=0 \\ -x_1+4x_2+x_3&=0 \end{aligned}$

9. $\begin{aligned} 2x_1+x_2-3x_3&=0 \\ 4x_1-x_2+x_3&=0 \end{aligned}$

10. $\begin{aligned} x_1+x_2&=0 \\ 2x_1+x_2&=0 \\ 3x_1+x_2&=0 \end{aligned}$

11. $\begin{aligned} x_1+x_2&=1 \\ 2x_1+x_2&=3 \\ 3x_1+x_2&=4 \end{aligned}$

12. $\begin{aligned} x_1+x_2+x_3+x_4&=4 \\ 2x_1-3x_2-x_3+4x_4&=7 \\ -2x_1+4x_2+x_3-2x_4&=1 \\ 5x_1-x_2+2x_3+x_4&=-1 \end{aligned}$

13. $\begin{aligned} x_1+x_2+x_3+x_4&=0 \\ 2x_1-3x_2-x_3+4x_4&=0 \\ -2x_1+4x_2+x_3-2x_4&=0 \\ 5x_1-x_2+2x_3+x_4&=0 \end{aligned}$

14. $\begin{aligned} x_1+x_2+x_3+x_4&=0 \\ 2x_1-3x_2-x_3+4x_4&=0 \\ -2x_1+4x_2+x_3-2x_4&=0 \end{aligned}$

En los Ejercicios del 15 al 22 realice los cálculos indicados.

15. $3\begin{pmatrix} -2 & 1 \\ 0 & 4 \\ 2 & 3 \end{pmatrix}$

16. $\begin{pmatrix} 1 & 0 & 3 \\ 2 & -1 & 6 \end{pmatrix}+\begin{pmatrix} 2 & 0 & 4 \\ -2 & 5 & 8 \end{pmatrix}$

17. $5\begin{pmatrix} 2 & 1 & 3 \\ -1 & 2 & 4 \\ -6 & 1 & 5 \end{pmatrix}-3\begin{pmatrix} -2 & 1 & 4 \\ 5 & 0 & 7 \\ 2 & -1 & 3 \end{pmatrix}$

18. $\begin{pmatrix} 2 & 3 \\ -1 & 4 \end{pmatrix}\begin{pmatrix} 5 & -1 \\ 2 & 7 \end{pmatrix}$

19. $\begin{pmatrix} 2 & 3 & 1 & 5 \\ 0 & 6 & 2 & 4 \end{pmatrix}\begin{pmatrix} 5 & 7 & 1 \\ 2 & 0 & 3 \\ 1 & 0 & 0 \\ 0 & 5 & 6 \end{pmatrix}$

20. $\begin{pmatrix} 2 & 3 & 5 \\ -1 & 6 & 4 \\ 1 & 0 & 6 \end{pmatrix}\begin{pmatrix} 0 & -1 & 2 \\ 3 & 1 & 2 \\ -7 & 3 & 5 \end{pmatrix}$

21. $\begin{pmatrix} 1 & 0 & 3 & -1 & 5 \\ 2 & 1 & 6 & 2 & 5 \end{pmatrix}\begin{pmatrix} 7 & 1 \\ 2 & 3 \\ -1 & 0 \\ 5 & 6 \\ 2 & 3 \end{pmatrix}$

22. $\begin{pmatrix} 1 & -1 & 2 \\ 3 & 5 & 6 \\ 2 & 4 & -1 \end{pmatrix}\begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}$

En los Ejercicios del 23 al 26 determine si la matriz dada está en forma escalonada por renglones (pero no en la forma escalonada por renglones reducida); en la forma escalonada por renglones reducida, o en ninguna de éstas.

23. $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{pmatrix}$

24. $\begin{pmatrix} 1 & 8 & 1 & 0 \\ 0 & 1 & 5 & -7 \\ 0 & 0 & 1 & 4 \end{pmatrix}$

25. $\begin{pmatrix} 1 & 0 \\ 0 & 3 \\ 0 & 0 \end{pmatrix}$

26. $\begin{pmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & 3 & 0 \end{pmatrix}$

En los Ejercicios 27 y 28, reduzca la matriz a forma escalonada por renglones y a la forma escalonada por renglones reducida.

27. $\begin{pmatrix} 2 & 8 & -2 \\ 1 & 0 & -6 \end{pmatrix}$

28. $\begin{pmatrix} 1 & -1 & 2 & 4 \\ -1 & 2 & 0 & 3 \\ 2 & 3 & -1 & 1 \end{pmatrix}$

En los Ejercicios del 29 al 33 calcule la forma escalonada por renglones y la inversa de la matriz dada (si existe).

29. $\begin{pmatrix} 2 & 3 \\ -1 & 4 \end{pmatrix}$

30. $\begin{pmatrix} -1 & 2 \\ 2 & -4 \end{pmatrix}$

31. $\begin{pmatrix} 1 & 2 & 0 \\ 2 & 1 & -1 \\ 3 & 1 & 1 \end{pmatrix}$

32. $\begin{pmatrix} -1 & 2 & 0 \\ 4 & 1 & -3 \\ 2 & 5 & -3 \end{pmatrix}$

33. $\begin{pmatrix} 2 & 0 & 4 \\ -1 & 3 & 1 \\ 0 & 1 & 2 \end{pmatrix}$

En los Ejercicios del 34 al 36, escriba primero el sistema en la forma $A\mathbf{x} = \mathbf{b}$, luego calcule A^{-1}, y finalmente, utilice la multiplicación de matrices para obtener el vector solución.

34. $\begin{aligned} x_1-3x_2&=4 \\ 2x_1+5x_2&=7 \end{aligned}$

35. $\begin{aligned} x_1+2x_2\phantom{{}-x_3}&=3 \\ 2x_1+x_2-x_3&=-1 \\ 3x_1+x_2+x_3&=7 \end{aligned}$

36. $\begin{aligned} 2x_1\phantom{{}+3x_2}+4x_3&=7 \\ -x_1+3x_2+x_3&=-4 \\ x_2+2x_3&=5 \end{aligned}$

*En los Ejercicios del 37 al 42 calcule la transpuesta de la matriz dada y determine si la matriz es simétrica o antisimétrica.**

37. $\begin{pmatrix} 2 & 3 & 1 \\ -1 & 0 & 2 \end{pmatrix}$

38. $\begin{pmatrix} 4 & 6 \\ 6 & 4 \end{pmatrix}$

39. $\begin{pmatrix} 2 & 3 & 1 \\ 3 & -6 & -5 \\ 1 & -5 & 9 \end{pmatrix}$

40. $\begin{pmatrix} 0 & 5 & 6 \\ -5 & 0 & 4 \\ -6 & -4 & 0 \end{pmatrix}$

41. $\begin{pmatrix} 1 & -1 & 4 & 6 \\ -1 & 2 & 5 & 7 \\ 4 & 5 & 3 & -8 \\ 6 & 7 & -8 & 9 \end{pmatrix}$

42. $\begin{pmatrix} 0 & 1 & -1 & 1 \\ -1 & 0 & 1 & -2 \\ 1 & 1 & 0 & 1 \\ 1 & -2 & -1 & 0 \end{pmatrix}$

* Del Problema 1.9.18 se tiene que: A es antisimétrica si $A^t = -A$.

En los Ejercicios del 43 *al* 47 *encuentre una matriz elemental* 3 × 3 *que realice las operaciones sobre renglones indicadas.*

43. $M_2(-2)$ **44.** $A_{2,1}(2)$ **45.** $A_{1,3}(-5)$ **46.** P_{31} **47.** $A_{3,2}(\frac{1}{5})$

En los Ejercicios del 48 *al* 50 *halle la inversa de la matriz elemental dada.*

48. $\begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}$ **49.** $\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ **50.** $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -\frac{1}{3} \end{pmatrix}$

En los Ejercicios 51 *y* 52 *escriba la matriz como el producto de matrices elementales.*

51. $\begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$ **52.** $\begin{pmatrix} 1 & 0 & 3 \\ 2 & 1 & -5 \\ 3 & 2 & 4 \end{pmatrix}$

En los Problemas 53 *y* 54 *escriba cada matriz como el producto de matrices elementales y una matriz triangular superior.*

53. $\begin{pmatrix} 2 & -1 \\ -4 & 2 \end{pmatrix}$ **54.** $\begin{pmatrix} 1 & -2 & 3 \\ 2 & 0 & 4 \\ 1 & 2 & 1 \end{pmatrix}$

55. a. Encuentre dos inversas derechas de $A = \begin{pmatrix} 1 & 3 & 2 \\ 2 & 7 & 5 \end{pmatrix}$.

b. Use el resultado de la parte (a) para obtener dos soluciones de $A\mathbf{x} = \begin{pmatrix} 4 \\ -2 \end{pmatrix}$.

56. Halle una inversa izquierda de $A = \begin{pmatrix} 2 & -1 \\ 4 & 2 \\ 1 & 5 \end{pmatrix}$.

CAPÍTULO 2

Determinantes

2.1 Definiciones

Sea $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ una matriz de 2×2. En la Sección 2.7 definimos el determinante de A como

$$\det A = a_{11}a_{22} - a_{12}a_{21} \tag{1}$$

Con frecuencia denotaremos det A como

$$|A| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \tag{2}$$

Mostraremos que A es invertible si y sólo si det $A \neq 0$. Como veremos, este importante teorema es válido para matrices de $n \times n$.

En este capítulo desarrollaremos algunas de las propiedades básicas de los determinantes, y veremos cómo pueden ser usadas para calcular inversas y resolver sistemas de n ecuaciones lineales en n incógnitas.

Definiremos *por inducción* el determinante de una matriz de $n \times n$. En otras palabras, usaremos nuestro conocimiento de un determinante de 2 × 2 para definir un determinante de 3 × 3; éste para definir el de 4 × 4 y así sucesivamente. Empezaremos por definir un determinante de 3 × 3.*

Definición 1 ***Determinante de 3 x 3.*** Sea $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$. Entonces

* Existen varias formas de definir un determinante y ésta es una de ellas. Es importante darse cuenta de que "det" es una función que asigna un *número* a una matriz *cuadrada*.

$$\det A = |A| = a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \quad (3)$$

Notemos el signo menos antes del segundo término del lado derecho de (3).

Ejemplo 1 Sea $A = \begin{pmatrix} 3 & 5 & 2 \\ 4 & 2 & 3 \\ -1 & 2 & 4 \end{pmatrix}$. Calcule $|A|$.

Solución

$$|A| = \begin{vmatrix} 3 & 5 & 2 \\ 4 & 2 & 3 \\ -1 & 2 & 4 \end{vmatrix} = 3\begin{vmatrix} 2 & 3 \\ 2 & 4 \end{vmatrix} - 5\begin{vmatrix} 4 & 3 \\ -1 & 4 \end{vmatrix} + 2\begin{vmatrix} 4 & 2 \\ -1 & 2 \end{vmatrix}$$
$$= 3\cdot 2 - 5\cdot 19 + 2\cdot 10 = -69$$

Ejemplo 2 Calcule $\begin{vmatrix} 2 & -3 & 5 \\ 1 & 0 & 4 \\ 3 & -3 & 9 \end{vmatrix}$.

Solución

$$\begin{vmatrix} 2 & -3 & 5 \\ 1 & 0 & 4 \\ 3 & -3 & 9 \end{vmatrix} = 2\begin{vmatrix} 0 & 4 \\ -3 & 9 \end{vmatrix} - (-3)\begin{vmatrix} 1 & 4 \\ 3 & 9 \end{vmatrix} + 5\begin{vmatrix} 1 & 0 \\ 3 & -3 \end{vmatrix}$$
$$= 2\cdot 12 + 3(-3) + 5(-3) = 0$$

Existe un método más sencillo para calcular determinantes de 3×3. De la Ecuación (3)

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31})$$

o sea
$$|A| = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{32}a_{23} \quad (4)$$

Escribimos A y junto a ella sus primeras dos columnas

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \begin{matrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{matrix}$$
$$- \quad - \quad - \qquad + \quad + \quad +$$

Luego se calculan los seis productos, poniendo signo menos antes de los productos con flechas que apuntan hacia arriba y efectuamos la suma. Con esto obtenemos la suma de la Ecuación 4.

Ejemplo 3 Calcule $\begin{vmatrix} 3 & 5 & 2 \\ 4 & 2 & 3 \\ -1 & 2 & 4 \end{vmatrix}$ usando este nuevo método

Solución Si escribimos $\begin{vmatrix} 3 & 5 & 2 \\ 4 & 2 & 3 \\ -1 & 2 & 4 \end{vmatrix} \begin{matrix} 3 & 5 \\ 4 & 2 \\ -1 & 2 \end{matrix}$ y multiplicamos como se indicó,

$$|A| = (3)(2)(4) + (5)(3)(-1) + (2)(4)(2) - (-1)(2)(2) - 2(3)(3) - (4)(4)(5)$$
$$= 24 - 15 + 16 + 4 - 18 - 80 = -69$$

Advertencia. El método expuesto anteriormente *no funciona* para determinantes de $n \times n$ si $n \neq 3$.

Antes de definir determinantes de $n\times n$, primero notemos que en la Ecuación (3), $\begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix}$ es la matriz que se obtiene al eliminar el primer renglón y la primera columna de A; $\begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{pmatrix}$ es la matriz que obtenemos si eliminamos el primer renglón y la segunda columna de A y $\begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}$ es la matriz que se obtiene si eliminamos el primer renglón y la tercera columna de A. Si denotamos estas tres matrices como M_{11}, M_{12} y M_{13}, respectivamente, y si $A_{11} = \det M_{11}$, $A_{12} = -\det M_{12}$ y $A_{13} = \det M_{13}$, entonces la Ecuación (3) puede ser escrita

$$\det A = |A| = a_{11}A_{11} + a_{12}A_{12} + a_{13}A_{13} \quad (5)$$

Definición 2 ***Menor.*** Sea A una matriz de $n \times n$ y sea M_{ij} la matriz de $(n-1) \times (n-1)$ que se obtiene de A, al eliminar el i-ésimo renglón y la j-ésima columna de A. M_{ij} se denomina *ij-ésimo menor* de A.

Ejemplo 4 Sea $A = \begin{pmatrix} 2 & -1 & 4 \\ 0 & 1 & 5 \\ 6 & 3 & -4 \end{pmatrix}$. Encuentre M_{13} y M_{32}.

Solución Eliminando el primer renglón y la tercera columna de A, obtenemos $M_{13} = \begin{pmatrix} 0 & 1 \\ 6 & 3 \end{pmatrix}$.

Análogamente, si eliminamos el tercer renglón y la segunda columna obtenemos $M_{32} = \begin{pmatrix} 2 & 4 \\ 0 & 5 \end{pmatrix}$.

Ejemplo 5 Sea $A = \begin{pmatrix} 1 & -3 & 5 & 6 \\ 2 & 4 & 0 & 3 \\ 1 & 5 & 9 & -2 \\ 4 & 0 & 2 & 7 \end{pmatrix}$. Encuentre M_{32} y M_{24}.

Solución Si eliminamos el tercer renglón y la segunda columna de A,

$M_{32} = \begin{pmatrix} 1 & 5 & 6 \\ 2 & 0 & 3 \\ 4 & 2 & 7 \end{pmatrix}$; análogamente, $M_{24} = \begin{pmatrix} 1 & -3 & 5 \\ 1 & 5 & 9 \\ 4 & 0 & 2 \end{pmatrix}$.

Definición 3 **Cofactor.** Sea A una matriz de $n \times n$. El *ij-ésimo cofactor* de A, denotado por A_{ij}, es

$$A_{ij} = (-1)^{i+j}|M_{ij}| \qquad (6)$$

Esto es, el ij-ésimo cofactor de A se obtiene tomando el determinante del ij-ésimo menor y multiplicándolo por $(-1)^{i+j}$. Notemos que

$$(-1)^{i+j} = \begin{cases} 1 & \text{si } i+j \text{ es par} \\ -1 & \text{si } i+j \text{ es impar} \end{cases}$$

Observación. La Definición 3 tiene sentido debido a que vamos a definir un determinante de $n \times n$ suponiendo que ya sabemos lo que es un determinante de $(n-1)\times(n-1)$.

Ejemplo 6 En el Ejemplo 5 tenemos que

$$A_{32} = (-1)^{3+2}|M_{32}| = -\begin{vmatrix} 1 & 5 & 6 \\ 2 & 0 & 3 \\ 4 & 2 & 7 \end{vmatrix} = -8$$

$$A_{24} = (-1)^{2+4}\begin{vmatrix} 1 & -3 & 5 \\ 1 & 5 & 9 \\ 4 & 0 & 2 \end{vmatrix} = -192$$

Consideremos ahora el caso de la matriz general de $n \times n$. Aquí

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \qquad (7)$$

Definición 4 **Determinante de n x n.** Sea A una matriz de $n \times n$. Entonces el determinante de A, escrito $\det A$, o bien $|A|$, está dado por

$$\det A = |A| = a_{11}A_{11} + a_{12}A_{12} + a_{13}A_{13} + \cdots + a_{1n}A_{1n} = \sum_{k=1}^{n} a_{1k}A_{1k} \qquad (8)$$

La expresión del lado derecho de (8) se conoce como *desarrollo por cofactores.*

Observación. En la Ecuación (8) quedó definido el determinante mediante el desarrollo por cofactores, para lo cual usamos las componentes del primer renglón de A. En la próxima sección (Teorema 2.2.1), veremos que se obtiene el mismo resultado si se desarrolla por cofactores en cualquier renglón o columna.

Ejemplo 7 Calcule $\det A$ si

$$A = \begin{pmatrix} 1 & 3 & 5 & 2 \\ 0 & -1 & 3 & 4 \\ 2 & 1 & 9 & 6 \\ 3 & 2 & 4 & 8 \end{pmatrix}$$

Solución

$$\begin{vmatrix} 1 & 3 & 5 & 2 \\ 0 & -1 & 3 & 4 \\ 2 & 1 & 9 & 6 \\ 3 & 2 & 4 & 8 \end{vmatrix} = a_{11}A_{11} + a_{12}A_{12} + a_{13}A_{13} + a_{14}A_{14}$$

$$= 1\begin{vmatrix} -1 & 3 & 4 \\ 1 & 9 & 6 \\ 2 & 4 & 8 \end{vmatrix} - 3\begin{vmatrix} 0 & 3 & 4 \\ 2 & 9 & 6 \\ 3 & 4 & 8 \end{vmatrix} + 5\begin{vmatrix} 0 & -1 & 4 \\ 2 & 1 & 6 \\ 3 & 2 & 8 \end{vmatrix} - 2\begin{vmatrix} 0 & -1 & 3 \\ 2 & 1 & 9 \\ 3 & 2 & 4 \end{vmatrix}$$

$$= 1(-92) - 3(-70) + 5(2) - 2(-16) = 160$$

Es claro que el cálculo del determinante de una matriz de $n \times n$ puede ser tedioso. Para calcular un determinante de 4×4, debemos evaluar cuatro determinantes de 3×3. Para calcular un determinante de 5×5, hay que evaluar cinco determinantes de 4×4, lo cual es lo mismo que calcular veinte determinantes de 3×3. Afortunadamente existen técnicas para simplificar enormemente estos cálculos. Algunos de estos métodos serán discutidos en la próxima sección. Existen, sin embargo, algunas matrices cuyos determinantes pueden ser fácilmente establecidos.

Definición 5 Una matriz cuadrada se llama *triangular superior* si todas sus componentes por debajo de la diagonal son cero. Será *triangular inferior* si todas sus compo-

nentes por encima de la diagonal son cero. Una matriz se llama *diagonal* si todos los elementos que no se encuentran en la diagonal son cero; así es que $A=(a_{ij})$ es triangular superior si $a_{ij}=0$ para $i>j$, triangular inferior si $a_{ij}=0$ para $i<j$ y diagonal si $a_{ij}=0$ para $i\neq j$. Notemos que una matriz diagonal es al mismo tiempo triangular superior e inferior.

Ejemplo 8 Las matrices $A=\begin{pmatrix}2&1&7\\0&2&-5\\0&0&1\end{pmatrix}$ y $B=\begin{pmatrix}-2&3&0&1\\0&0&2&4\\0&0&1&3\\0&0&0&-2\end{pmatrix}$ son triangulares superiores; $C=\begin{pmatrix}5&0&0\\2&3&0\\-1&2&4\end{pmatrix}$ y $D=\begin{pmatrix}0&0\\1&0\end{pmatrix}$ son triangulares inferiores; I y $E=\begin{pmatrix}2&0&0\\0&-7&0\\0&0&-4\end{pmatrix}$ son diagonales.

Ejemplo 9 Sea

$$A=\begin{pmatrix}a_{11}&0&0&0\\a_{21}&a_{22}&0&0\\a_{31}&a_{32}&a_{33}&0\\a_{41}&a_{42}&a_{43}&a_{44}\end{pmatrix}$$

triangular inferior. Calcule $\det A$.

Solución

$$\det A=a_{11}A_{11}+0A_{12}+0A_{13}+0A_{14}=a_{11}A_{11}$$

$$=a_{11}\begin{vmatrix}a_{22}&0&0\\a_{32}&a_{33}&0\\a_{42}&a_{43}&a_{44}\end{vmatrix}$$

$$=a_{11}a_{22}\begin{vmatrix}a_{33}&0\\a_{43}&a_{44}\end{vmatrix}$$

$$=a_{11}a_{22}a_{33}a_{44}$$

El Ejemplo 9 puede ser generalizado fácilmente para demostrar el siguiente teorema.

Teorema 1 Sea $A=(a_{ij})$ una matriz de $n\times n$ triangular superior* o inferior. Entonces

* La demostración para el caso de la matriz triangular superior es más complicada a esta altura, pero será exactamente la misma una vez que sepamos que $\det A$ puede ser evaluado desarrollando por cualquier columna (Teorema 2.2.1).

$$\det A=a_{11}a_{22}a_{33}\cdots a_{nn} \qquad (9)$$

Es decir: *El determinante de una matriz triangular es igual al producto de sus componentes diagonales.*

Ejemplo 10 Los determinantes de las seis matrices del Ejemplo 8 son: $|A|=2\cdot 2\cdot 1=4$; $|B|=(-2)(0)(1)(-2)=0$; $|C|=5\cdot 3\cdot 4=60$; $|D|=0$; $|I|=1$; $|E|=(2)(-7)(-4)=56$.

Problemas 2.1

En los Problemas del 1 *al* 10 *calcule el determinante.*

1. $\begin{vmatrix}1&0&3\\0&1&4\\2&1&0\end{vmatrix}$ **2.** $\begin{vmatrix}-1&1&0\\2&1&4\\1&5&6\end{vmatrix}$ **3.** $\begin{vmatrix}3&-1&4\\6&3&5\\2&-1&6\end{vmatrix}$

4. $\begin{vmatrix}-1&0&6\\0&2&4\\1&2&-3\end{vmatrix}$ **5.** $\begin{vmatrix}-2&3&1\\4&6&5\\0&2&1\end{vmatrix}$ **6.** $\begin{vmatrix}5&-2&1\\6&0&3\\-2&1&4\end{vmatrix}$

7. $\begin{vmatrix}2&0&3&1\\0&1&4&2\\0&0&1&5\\1&2&3&0\end{vmatrix}$ **8.** $\begin{vmatrix}-3&0&0&0\\-4&7&0&0\\5&8&-1&0\\2&3&0&6\end{vmatrix}$ **9.** $\begin{vmatrix}-2&0&0&7\\1&2&-1&4\\3&0&-1&5\\4&2&3&0\end{vmatrix}$

10. $\begin{vmatrix}2&3&-1&4&5\\0&1&7&8&2\\0&0&4&-1&5\\0&0&0&-2&8\\0&0&0&0&6\end{vmatrix}$

11. Demuestre que si A y B son matrices diagonales de $n\times n$, entonces $\det AB=\det A\,\det B$.

★ **12.** Demuestre que si A y B son matrices triangulares inferiores, entonces $\det AB=\det A\,\det B$.

13. Pruebe que, en general, no es cierto que $\det(A+B)=\det A+\det B$.

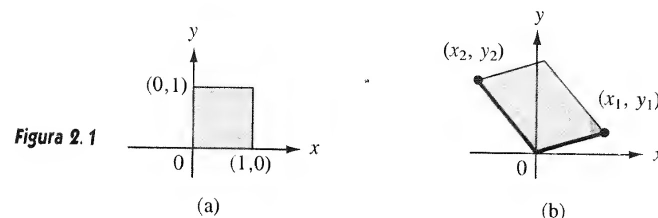
14. Pruebe que si A es triangular, entonces $\det A\neq 0$ si y sólo si todas las componentes diagonales de A son diferentes de cero.

15. Demuestre el Teorema 1 para una matriz triangular superior.

★ **16.** Decimos que los vectores $\begin{pmatrix}0\\1\end{pmatrix}$ y $\begin{pmatrix}1\\0\end{pmatrix}$ *generan el área* 1 en el plano, ya que si construimos un cuadrado con tres de sus vértices en (0, 0), (1, 0) y (0, 1), vemos que el

área es 1. (Figura 2.1*a*.) Más generalmente, si $\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$ y $\begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$ son dos vectores linealmente independientes, de dos componentes, entonces generan un área que se define como el área del paralelogramo con tres de sus cuatro vértices en (0, 0), (x_1, y_1) y (x_2, y_2). (Figura 2.1*b*.)

Figura 2.1



Sea A una matriz de 2×2. Si k denota el área generada por $\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$ y $\begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$, donde $\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = A\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ y $\begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = A\begin{pmatrix} 0 \\ 1 \end{pmatrix}$, demuestre que $k = |\det A|$.

★★ **17.** Sean $\mathbf{u}_1$ y $\mathbf{u}_2$ dos vectores de dos componentes y sean $\mathbf{v}_1 = A\mathbf{u}_1$ y $\mathbf{v}_2 = A\mathbf{u}_2$. Demuestre que (área generada por $\mathbf{v}_1$ y $\mathbf{v}_2$) = (área generada por $\mathbf{u}_1$ y $\mathbf{u}_2$) $|\det A|$.

Esto proporciona una interpretación geométrica del determinante.

2.2 Propiedades de los determinantes

Los determinantes tienen muchas propiedades que pueden facilitar los cálculos. Empezaremos a describir estas propiedades estableciendo un teorema, del cual deduciremos lo demás. La demostración de este teorema es difícil y se pospondrá para la próxima sección.

Teorema 1 **Teorema básico.** Sea

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

una matriz de $n\times n$. Entonces

$$\det A = a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in} = \sum_{k=1}^{n} a_{ik}A_{ik} \qquad (1)$$

para $i = 1, 2, \ldots, n$. Es decir, podemos calcular $\det A$ desarrollando por cofactores en *cualquier* renglón de A. Más aún:

$$\det A = a_{1j}A_{1j} + a_{2j}A_{2j} + \cdots + a_{nj}A_{nj} = \sum_{k=1}^{n} a_{kj}A_{kj} \qquad (2)$$

Dado que la j-ésima columna de A es $\begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{pmatrix}$, la Ecuación (2) nos indica que podemos calcular el $\det A$ desarrollando por cofactores en cualquier columna de A.

Ejemplo 1 Para $A = \begin{pmatrix} 3 & 5 & 2 \\ 4 & 2 & 3 \\ -1 & 2 & 4 \end{pmatrix}$, vimos en el Ejemplo 2.1.1 que $\det A = -69$. Si desarrollamos en el segundo renglón resulta que

$$\begin{aligned} \det A &= 4A_{21} + 2A_{22} + 3A_{23} \\ &= 4(-1)^{2+1}\begin{vmatrix} 5 & 2 \\ 2 & 4 \end{vmatrix} + 2(-1)^{2+2}\begin{vmatrix} 3 & 2 \\ -1 & 4 \end{vmatrix} + 3(-1)^{2+3}\begin{vmatrix} 3 & 5 \\ -1 & 2 \end{vmatrix} \\ &= -4(16) + 2(14) - 3(11) = -69 \end{aligned}$$

Análogamente, si se desarrolla en la tercera columna, por ejemplo,

$$\begin{aligned} \det A &= 2A_{13} + 3A_{23} + 4A_{33} \\ &= 2(-1)^{1+3}\begin{vmatrix} 4 & 2 \\ -1 & 2 \end{vmatrix} + 3(-1)^{2+3}\begin{vmatrix} 3 & 5 \\ -1 & 2 \end{vmatrix} + 4(-1)^{3+3}\begin{vmatrix} 3 & 5 \\ 4 & 2 \end{vmatrix} \\ &= 2(10) - 3(11) + 4(-14) = -69 \end{aligned}$$

Inténtese verificar que obtenemos el mismo resultado si desarrollamos en el tercer renglón o la primera o la segunda columna.

Enunciaremos y demostraremos algunas propiedades adicionales de los determinantes. En cada caso vamos a suponer que A es una matriz de $n \times n$.* Se verá que estas propiedades pueden ser usadas para reducir enormemente el trabajo de calcular un determinante.

Propiedad 1 Si cualquier renglón o columna de A es el vector cero, entonces $\det A = 0$.

Demostración Supongamos que el i-ésimo renglón de A contiene únicamente ceros. Es decir, $a_{ij} = 0$ para $j = 1, 2, \ldots, n$. Entonces $\det A = a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in} = 0 + 0 + \cdots + 0 = 0$. La misma demostración funciona si la j-ésima columna es el vector cero. ■

* Las demostraciones de estas propiedades están dadas en términos de los renglones de una matriz. Si usamos el Teorema 1, las mismas propiedades pueden ser demostradas para las columnas.

Ejemplo 2 Es fácil verificar que

$$\begin{vmatrix} 2 & 3 & 5 \\ 0 & 0 & 0 \\ 1 & -2 & 4 \end{vmatrix} = 0 \quad \text{y} \quad \begin{vmatrix} -1 & 3 & 0 & 1 \\ 4 & 2 & 0 & 5 \\ -1 & 6 & 0 & 4 \\ 2 & 1 & 0 & 1 \end{vmatrix} = 0$$

Propiedad 2 Si el i-ésimo renglón o la j-ésima columna de A se multiplican por la constante c, entonces det A se multiplica por c. Es decir, si llamamos a esta nueva matriz B, entonces

$$|B| = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ ca_{i1} & ca_{i2} & \cdots & ca_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = c \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = c|A| \qquad (3)$$

Demostración Para demostrar (3) expandimos en el i-ésimo renglón de A:

$$\det B = ca_{i1}A_{i1} + ca_{i2}A_{i2} + \cdots ca_{in}A_{in}$$
$$= c(a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in}) = c \det A$$

Una demostración análoga funciona para las columnas. ■

Ejemplo 3 Sea $A = \begin{pmatrix} 1 & -1 & 2 \\ 3 & 1 & 4 \\ 0 & -2 & 5 \end{pmatrix}$. Entonces det $A = 16$. Si multiplicamos el segundo renglón por 4, tenemos $B = \begin{pmatrix} 1 & -1 & 2 \\ 12 & 4 & 16 \\ 0 & -2 & 5 \end{pmatrix}$ y det $B = 64 = 4$ det A. Si se multiplica la tercera columna por -3, obtenemos $C = \begin{pmatrix} 1 & -1 & -6 \\ 3 & 1 & -12 \\ 0 & -2 & -15 \end{pmatrix}$ y det $C = -48 = -3$ det A.

Observación. Si usamos la Propiedad 2 podemos demostrar (Problema 28) el siguiente resultado interesante: Para cualquier escalar α y cualquier matriz A de $n \times n$, $\det \alpha A = \alpha^n \det A$.

Propiedad 3 Sean

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nj} & \cdots & a_{nn} \end{pmatrix}, \quad B = \begin{pmatrix} a_{11} & a_{12} & \cdots & \alpha_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & \alpha_{2j} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & \alpha_{nj} & \cdots & a_{nn} \end{pmatrix},$$

y

$$C = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1j}+\alpha_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j}+\alpha_{2j} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nj}+\alpha_{nj} & \cdots & a_{nn} \end{pmatrix}$$

Entonces

$$\det C = \det A + \det B \qquad (4)$$

En otras palabras, supongamos que A, B y C son idénticas excepto por la j-ésima columna y que la j-ésima columna de C es la suma de las j-ésimas columnas de A y B. Entonces det C = det A + det B. Esto mismo es válido para renglones.

Demostración Se desarrolla det C en la j-ésima columna para obtener

$$\det C = (a_{1j} + \alpha_{1j})A_{1j} + (a_{2j} + \alpha_{2j})A_{2j} + \cdots + (a_{nj} + \alpha_{nj})A_{nj}$$
$$= (a_{1j}A_{1j} + a_{2j}A_{2j} + \cdots + a_{nj}A_{nj})$$
$$+ (\alpha_{1j}A_{1j} + \alpha_{2j}A_{2j} + \cdots + \alpha_{nj}A_{nj}) = \det A + \det B \quad ■$$

Ejemplo 4 Sea $A = \begin{pmatrix} 1 & -1 & 2 \\ 3 & 1 & 4 \\ 0 & -2 & 5 \end{pmatrix}$, $B = \begin{pmatrix} 1 & -6 & 2 \\ 3 & 2 & 4 \\ 0 & 4 & 5 \end{pmatrix}$, y $C = \begin{pmatrix} 1 & -1-6 & 2 \\ 3 & 1+2 & 4 \\ 0 & -2+4 & 5 \end{pmatrix} = \begin{pmatrix} 1 & -7 & 2 \\ 3 & 3 & 4 \\ 0 & 2 & 5 \end{pmatrix}$. Entonces det $A = 16$, det $B = 108$ y det $C = 124 =$ det $A +$ det B.

Propiedad 4 Si se intercambian dos renglones (o columnas) cualesquiera de A, es como si se multiplicara det A por -1.

Demostración Haremos la demostración para renglones y vamos a suponer primero que se intercambian dos renglones adyacentes. Es decir, vamos a suponer que se intercambian el i-ésimo y el $(i + 1)$-ésimo renglón. Sean

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & & \vdots\\ a_{i1} & a_{i2} & \cdots & a_{in}\\ a_{i+1,1} & a_{i+1,2} & \cdots & a_{i+1,n}\\ \vdots & \vdots & & \vdots\\ a_{n1} & a_{n2} & & a_{nn}\end{pmatrix} \quad \text{y} \quad B=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & & \vdots\\ a_{i+1,1} & a_{i+1,2} & \cdots & a_{i+1,n}\\ a_{i1} & a_{i2} & \cdots & a_{in}\\ \vdots & \vdots & & \vdots\\ a_{n1} & a_{n2} & \cdots & a_{nn}\end{pmatrix}$$

Entonces, si se desarrolla $\det A$ en su i-ésimo renglón y $\det B$ en su $(i+1)$-ésimo renglón,

$$\begin{aligned}\det A &= a_{i1}A_{i1}+a_{i2}A_{i2}+\cdots+a_{in}A_{in}\\ \det B &= a_{i1}B_{i+1,1}+a_{i2}B_{i+1,2}+\cdots+a_{in}B_{i+1,n}\end{aligned} \tag{5}$$

Aquí $A_{ij}=(-1)^{i+j}|M_{ij}|$, donde M_{ij} se obtiene eliminando el i-ésimo renglón y la j-ésima columna de A. Notemos ahora que si eliminamos el $(i+1)$-ésimo renglón y la j-ésima columna de B, obtenemos el mismo M_{ij}. Así

$$B_{i+1,j}=(-1)^{i+1+j}|M_{ij}|=-(-1)^{i+j}|M_{ij}|=-A_{ij}$$

por lo tanto, de las Ecuaciones (5), $\det B=-\det A$.

Ahora supongamos que $i<j$ y que se intercambian el i-ésimo y el j-ésimo renglones. Esto se logra intercambiando renglones adyacentes varias veces. Necesitaremos $j-i$ intercambios para mover el j-ésimo renglón al i-ésimo renglón. Entonces el i-ésimo renglón se habrá movido al $(i+1)$-ésimo renglón, por lo cual se necesitarán $j-i-1$ intercambios adicionales para mover el renglón i al j-ésimo renglón. Para ilustrar, intercambiemos los renglones 2 y 6:*

$$\begin{array}{cccccccccccccccc} 1 & & 1 & & 1 & & 1 & & 1 & & 1 & & 1 & & 1\\ 2 & & 2 & & 2 & & 2 & & 6 & & 6 & & 6 & & 6\\ 3 & \to & 3 & \to & 3 & \to & 6 & \to & 2 & \to & 3 & \to & 3 & \to & 3\\ 4 & & 4 & & 6 & & 3 & & 3 & & 2 & & 4 & & 4\\ 5 & & 6 & & 4 & & 4 & & 4 & & 4 & & 2 & & 5\\ 6 & & 5 & & 5 & & 5 & & 5 & & 5 & & 5 & & 2\\ 7 & & 7 & & 7 & & 7 & & 7 & & 7 & & 7 & & 7 \end{array}$$

$6-2=4$ intercambios para mover el 6 a la posición del 2.

$6-2-1=3$ intercambios para mover el 2 a la posición del 6.

Finalmente, el número total de intercambios de renglones adyacentes es $(j-i)+(j-i-1)=2j-2i-1$, el cual es impar. Así, el $\det A$ se multiplica por -1 un número impar de veces, lo cual es lo que necesitábamos demostrar. ■

* Note que aquí todos los números se refieren a renglones.

Ejemplo 5 Sea $A=\begin{pmatrix}1 & -1 & 2\\ 3 & 1 & 4\\ 0 & -2 & 5\end{pmatrix}$. Si intercambiamos el primero y el tercer renglones, obtenemos $B=\begin{pmatrix}0 & -2 & 5\\ 3 & 1 & 4\\ 1 & -1 & 2\end{pmatrix}$. Si intercambiamos la primera y la segunda columnas de A, obtenemos $C=\begin{pmatrix}-1 & 1 & 2\\ 1 & 3 & 4\\ -2 & 0 & 5\end{pmatrix}$. Entonces, por cálculos directos, $\det A=16$ y $\det B=\det C=-16$.

Propiedad 5 Si A tiene dos renglones o columnas iguales, entonces $\det A=0$.

Demostración Supongamos que el i-ésimo y el j-ésimo renglones de A son iguales. Si intercambiamos estos renglones obtenemos una matriz B con la propiedad de que $\det B=-\det A$ (de la Propiedad 4). Pero puesto que renglón i = renglón j, si los intercambiamos se obtiene la misma matriz. Así, $A=B$ y $\det A=\det B=-\det A$. De donde $2\det A=0$, lo cual sólo puede pasar si $\det A=0$.

Ejemplo 6 Por cálculos directos podemos verificar que para $A=\begin{pmatrix}1 & -1 & 2\\ 5 & 7 & 3\\ 1 & -1 & 2\end{pmatrix}$ [dos renglones iguales] y $B=\begin{pmatrix}5 & 2 & 2\\ 3 & -1 & -1\\ -2 & 4 & 4\end{pmatrix}$ [dos columnas iguales], se tiene que $\det A=\det B=0$.

Propiedad 6 Si un renglón (columna) de A es un múltiplo constante de otro renglón (columna), entonces $\det A=0$.

Demostración Sea $(a_{j1}, a_{j2}, \ldots, a_{jn})=c(a_{i1}, a_{i2}, \ldots, a_{in})$. Entonces de la Propiedad 2,

$$\det A = c\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & & \vdots\\ a_{i1} & a_{i2} & \cdots & a_{in}\\ \vdots & \vdots & & \vdots\\ a_{i1} & a_{i2} & \cdots & a_{in}\\ \vdots & \vdots & & \vdots\\ a_{n1} & a_{n2} & \cdots & a_{nn}\end{vmatrix} = 0 \quad \text{(de la Propiedad 5)}$$

(j-ésimo renglón → $a_{i1}\ a_{i2}\ \cdots\ a_{in}$) ■

Ejemplo 7 $\begin{vmatrix} 2 & -3 & 5 \\ 1 & 7 & 2 \\ -4 & 6 & -10 \end{vmatrix} = 0$ puesto que el tercer renglón es el primer renglón multiplicado por -2.

Ejemplo 8 $\begin{vmatrix} 2 & 4 & 1 & 12 \\ -1 & 1 & 0 & 3 \\ 0 & -1 & 9 & -3 \\ 7 & 3 & 6 & 9 \end{vmatrix} = 0$ puesto que la cuarta columna es tres veces la segunda columna.

Propiedad 7 Si un múltiplo de un renglón (columna) de A se suma a otro renglón (columna) de A, el determinante no cambiará.

Demostración Sea B la matriz obtenida sumando c veces el i-ésimo renglón de A al j-ésimo renglón de A. Entonces

$$\det B = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{j1}+ca_{i1} & a_{j2}+ca_{i2} & \cdots & a_{jn}+ca_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

$$\text{(por la Propiedad 3)} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ ca_{i1} & ca_{i2} & \cdots & ca_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

$= \det A + 0 = \det A$ (el cero se debe a la Propiedad 6) ■

Ejemplo 9 Sea $A = \begin{pmatrix} 1 & -1 & 2 \\ 3 & 1 & 4 \\ 0 & -2 & 5 \end{pmatrix}$ Entonces $\det A = 16$. Si multiplicamos el tercer renglón por 4 y lo sumamos al segundo renglón, obtenemos una nueva matriz B dada por

$$B = \begin{pmatrix} 1 & -1 & 2 \\ 3+4(0) & 1+4(-2) & 4+5(4) \\ 0 & -2 & 5 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 2 \\ 3 & -7 & 24 \\ 0 & -2 & 5 \end{pmatrix}$$

y así $\det B = 16 = \det A$.

Las propiedades que preceden hacen mucho más fácil el cálculo de determinantes de órdenes elevados. Simplemente "reducimos por renglón" el determinante, usando la Propiedad 7, hasta que el determinante se reduzca a una forma más fácil de evaluar. Lo más común es usar la Propiedad 7 repetidamente hasta que (*i*) el nuevo determinante tenga un renglón (columna) de ceros, o un renglón (columna) sea un múltiplo de otro renglón (columna), en cuyo caso el determinante es cero, o (*ii*) la nueva matriz sea triangular de modo que su determinante sea el producto de sus elementos diagonales.

Ejemplo 10 Calcule

$$|A| = \begin{vmatrix} 1 & 3 & 5 & 2 \\ 0 & -1 & 3 & 4 \\ 2 & 1 & 9 & 6 \\ 3 & 2 & 4 & 8 \end{vmatrix}$$

Solución (Vea el Ejemplo 2.1.7.)

Tenemos ya un cero en la primera columna, así que es más simple reducir a cero otros elementos en la primera columna. Luego se continuará reduciendo, tratando de obtener una matriz triangular:

Multiplicamos el primer renglón por -2 y lo sumamos al tercer renglón y se multiplica el primer renglón por -3 y se le suma al cuarto renglón.

$$|A| = \begin{vmatrix} 1 & 3 & 5 & 2 \\ 0 & -1 & 3 & 4 \\ 0 & -5 & -1 & 2 \\ 0 & -7 & -11 & 2 \end{vmatrix}$$

Multiplicamos el segundo renglón por -5 y -7 y lo sumamos al tercer y cuarto renglones, respectivamente.

$$= \begin{vmatrix} 1 & 3 & 5 & 2 \\ 0 & -1 & 3 & 4 \\ 0 & 0 & -16 & -18 \\ 0 & 0 & -32 & -26 \end{vmatrix}$$

Factorizamos -16 del tercer renglón (usando la Propiedad 2).

$$= -16 \begin{vmatrix} 1 & 3 & 5 & 2 \\ 0 & -1 & 3 & 4 \\ 0 & 0 & 1 & \frac{9}{8} \\ 0 & 0 & -32 & -26 \end{vmatrix}$$

Multiplicamos el tercer renglón por 32 y se le suma al cuarto renglón.

$$= -16\begin{vmatrix} 1 & 3 & 5 & 2 \\ 0 & -1 & 3 & 4 \\ 0 & 0 & 1 & \frac{9}{8} \\ 0 & 0 & 0 & 10 \end{vmatrix}$$

Ahora queda una matriz triangular superior y $|A| = -16(1)(-1)(1)(10) = (-16)(-10) = 160$.

Ejemplo 11 Calcule

$$|A| = \begin{vmatrix} -2 & 1 & 0 & 4 \\ 3 & -1 & 5 & 2 \\ -2 & 7 & 3 & 1 \\ 3 & -7 & 2 & 5 \end{vmatrix}$$

Solución En este caso existen muchas maneras de proceder y no es claro cuál camino nos llevará más rápidamente a la respuesta. Sin embargo, puesto que ya existe un cero en el primer renglón, empezaremos la reducción en ese renglón.

Multiplicamos la segunda columna por 2 y -4 y la sumamos a la primera y cuarta columnas, respectivamente.

$$|A| = \begin{vmatrix} 0 & 1 & 0 & 0 \\ 1 & -1 & 5 & 6 \\ 12 & 7 & 3 & -27 \\ -11 & -7 & 2 & 33 \end{vmatrix}$$

Intercambiamos las primeras dos columnas.

$$= -\begin{vmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 5 & 6 \\ 7 & 12 & 3 & -27 \\ -7 & -11 & 2 & 33 \end{vmatrix}$$

Multiplicamos la segunda columna por -5 y -6, y la sumamos a la tercera y cuarta columnas, respectivamente.

$$= -\begin{vmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 7 & 12 & -57 & -99 \\ -7 & -11 & 57 & 99 \end{vmatrix}$$

Puesto que la cuarta columna es ahora un múltiplo de la tercera (columna $4 = \frac{99}{57} \times$ columna 3), vemos que $|A| = 0$.

Ejemplo 2 Calcule

$$|A| = \begin{vmatrix} 1 & -2 & 3 & -5 & 7 \\ 2 & 0 & -1 & -5 & 6 \\ 4 & 7 & 3 & -9 & 4 \\ 3 & 1 & -2 & -2 & 3 \\ -5 & -1 & 3 & 7 & -9 \end{vmatrix}$$

Solución Si sumamos primero el renglón 2 y después el renglón 4 al renglón 5, obtenemos

$$|A| = \begin{vmatrix} 1 & -2 & 3 & 5 & 7 \\ 2 & 0 & -1 & -5 & 6 \\ 4 & 7 & 3 & -9 & 4 \\ 3 & 1 & -2 & -2 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{vmatrix} = 0 \quad \text{(de la Propiedad 1)}$$

Con este ejemplo se ilustra que si observamos un poco antes de hacer los cálculos, podemos simplificar el trabajo considerablemente.

Hay otras tres propiedades de los determinantes que nos serán muy útiles.

Teorema 2 Sea A una matriz de $n \times n$. Entonces

$$a_{i1}A_{j1} + a_{i2}A_{j2} + \cdots + a_{in}A_{jn} = 0 \quad \text{si } i \neq j \qquad (6)$$

Nota. Del Teorema 1 vemos que la suma en la Ecuación (6) es igual a $\det A$ si $i = j$.

Demostración Sea

$$B = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{matrix} \\ \\ \\ \\ \\ \leftarrow j\text{-ésimo renglón} \\ \\ \\ \end{matrix}$$

Entonces, como dos renglones de B son iguales, det $B = 0$. Pero $B = A$ excepto en el j-ésimo renglón. De esta forma, si calculamos el det B desarrollando en el j-ésimo renglón de B, obtendremos la suma de la Ecuación (6) y el teorema queda demostrado. Notemos que cuando se calculan los cofactores de B desarrollando en el j-ésimo renglón, éste desaparece. Así es que $B_{jk} = A_{jk}$ para $k = 1, 2, \ldots, n$. ■

Teorema 3 Sea A una matriz de $n \times n$. Entonces

$$\det A = \det A^t \tag{7}$$

Demostración Para esta demostración se requiere inducción matemática. Si no está familiarizado con este importante método de demostración, consulte el Apéndice 1. Demostraremos primero el teorema para el caso $n = 2$. Si

$$|A| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

entonces

$$|A^t| = \begin{vmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{21}a_{12} = |A|$$

y así el teorema es válido para $n = 2$. Ahora supondremos que el teorema es válido para las matrices de $(n-1) \times (n-1)$ y se lo demostrará para las matrices de $n \times n$. Lo anterior prueba el teorema. Sea $B = A^t$. Entonces

$$|A| = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \quad \text{y} \quad |A^t| = |B| = \begin{vmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{vmatrix}$$

Desarrollamos $|A|$ en el primer renglón y $|B|$ en la primera columna. Con esto

$$|A| = a_{11}A_{11} + a_{12}A_{12} + \cdots + a_{1n}A_{1n}$$

$$|B| = a_{11}B_{11} + a_{12}B_{21} + \cdots + a_{1n}B_{n1}$$

Es necesario demostrar que $A_{1k} = B_{k1}$ para $k = 1, 2, \ldots, n$. Pero $A_{1k} = (-1)^{1+k}|M_{1k}|$ y $B_{k1} = (-1)^{k+1}|N_{k1}|$, donde M_{1k} es el $1k$-ésimo menor de A y N_{k1} es el $k1$-ésimo menor de B. Entonces

$$|M_{1k}| = \begin{vmatrix} a_{21} & a_{22} & \cdots & a_{2,k-1} & a_{2,k+1} & \cdots & a_{2n} \\ a_{31} & a_{32} & \cdots & a_{3,k-1} & a_{3,k+1} & \cdots & a_{3n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{n,k-1} & a_{n,k+1} & \cdots & a_{nn} \end{vmatrix}$$

y

$$|N_{k1}| = \begin{vmatrix} a_{21} & a_{31} & \cdots & a_{n1} \\ a_{22} & a_{32} & \cdots & a_{n2} \\ \vdots & \vdots & & \vdots \\ a_{2,k-1} & a_{3,k-1} & \cdots & a_{n,k-1} \\ a_{2,k+1} & a_{3,k+1} & \cdots & a_{n,k+1} \\ \vdots & \vdots & & \vdots \\ a_{2n} & a_{3n} & \cdots & a_{nn} \end{vmatrix}$$

Claramente $M_{1k} = N_{k1}^t$, y puesto que ambas son matrices de $(n-1) \times (n-1)$, la hipótesis de inducción nos dice que $|M_{1k}| = |N_{k1}|$. Así $A_{1k} = B_{k1}$ y la demostración está completa. ■

Ejemplo 13 Sea $A = \begin{pmatrix} 1 & -1 & 2 \\ 3 & 1 & 4 \\ 0 & -2 & 5 \end{pmatrix}$. Entonces $A^t = \begin{pmatrix} 1 & 3 & 0 \\ -1 & 1 & -2 \\ 2 & 4 & 5 \end{pmatrix}$ y es fácil verificar que $|A| = |A^t| = 16$.

Teorema 4 Sean A y B matrices de $n \times n$. Entonces

$$\det AB = \det A \det B \tag{8}$$

Es decir: *El determinante del producto es el producto de los determinantes.*

Demostración La demostración con matrices elementales se da en la Sección 2.3. En el Problema 38 se pide verificar el resultado en el caso de 2×2. ■

Ejemplo 14 Verifique la Ecuación (8) para $A = \begin{pmatrix} 1 & -1 & 2 \\ 3 & 1 & 4 \\ 0 & -2 & 5 \end{pmatrix}$ y $B = \begin{pmatrix} 1 & -2 & 3 \\ 0 & -1 & 4 \\ 2 & 0 & -2 \end{pmatrix}$.

Solución det $A = 16$ y det $B = -8$. Calculamos

$$AB = \begin{pmatrix} 1 & -1 & 2 \\ 3 & 1 & 4 \\ 0 & -2 & 5 \end{pmatrix}\begin{pmatrix} 1 & -2 & 3 \\ 0 & -1 & 4 \\ 2 & 0 & -2 \end{pmatrix} = \begin{pmatrix} 5 & -1 & -5 \\ 11 & -7 & 5 \\ 10 & 2 & -18 \end{pmatrix}$$

y así det $AB = -128 = (16)(-8) =$ det A det B.

Problemas 2.2

En los Problemas del 1 al 20 evalúe el determinante usando los métodos de esta sección.

1. $\begin{vmatrix} 3 & -5 \\ 2 & 6 \end{vmatrix}$

2. $\begin{vmatrix} 4 & 1 \\ 0 & -3 \end{vmatrix}$

3. $\begin{vmatrix} -1 & 0 & 2 \\ 3 & 1 & 4 \\ 2 & 0 & -6 \end{vmatrix}$

4. $\begin{vmatrix} 2 & 1 & -1 \\ 3 & -2 & 0 \\ 5 & 1 & 6 \end{vmatrix}$

5. $\begin{vmatrix} -3 & 2 & 4 \\ 1 & -1 & 2 \\ -1 & 4 & 0 \end{vmatrix}$

6. $\begin{vmatrix} 0 & -2 & 3 \\ 1 & 2 & -3 \\ 4 & 0 & 5 \end{vmatrix}$

7. $\begin{vmatrix} -2 & 3 & 6 \\ 4 & 1 & 8 \\ -2 & 0 & 0 \end{vmatrix}$

8. $\begin{vmatrix} 2 & -1 & 3 \\ 4 & 0 & 6 \\ 5 & -2 & 3 \end{vmatrix}$

9. $\begin{vmatrix} 1 & -1 & 2 & 4 \\ 0 & -3 & 5 & 6 \\ 1 & 4 & 0 & 3 \\ 0 & 5 & -6 & 7 \end{vmatrix}$

10. $\begin{vmatrix} 2 & -3 & 1 & 4 \\ 0 & -2 & 0 & 0 \\ 3 & 7 & -1 & 2 \\ 4 & 1 & -3 & 8 \end{vmatrix}$

11. $\begin{vmatrix} 1 & 1 & -1 & 0 \\ -3 & 4 & 6 & 0 \\ 2 & 5 & -1 & 3 \\ 4 & 0 & 3 & 0 \end{vmatrix}$

12. $\begin{vmatrix} 3 & -1 & 2 & 1 \\ 4 & 3 & 1 & -2 \\ -1 & 0 & 2 & 3 \\ 6 & 2 & 5 & 2 \end{vmatrix}$

13. $\begin{vmatrix} 2 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 4 \end{vmatrix}$

14. $\begin{vmatrix} 0 & a & 0 & 0 \\ b & 0 & 0 & 0 \\ 0 & 0 & 0 & c \\ 0 & 0 & d & 0 \end{vmatrix}$

15. $\begin{vmatrix} 1 & 2 & 0 & 0 \\ 3 & -2 & 0 & 0 \\ 0 & 0 & 1 & -5 \\ 0 & 0 & 7 & 2 \end{vmatrix}$

16. $\begin{vmatrix} a & b & 0 & 0 \\ c & d & 0 & 0 \\ 0 & 0 & a & -b \\ 0 & 0 & c & d \end{vmatrix}$

17. $\begin{vmatrix} 2 & -1 & 0 & 4 & 1 \\ 3 & 1 & -1 & 2 & 0 \\ 3 & 2 & -2 & 5 & 1 \\ 0 & 0 & 4 & -1 & 6 \\ 3 & 2 & 1 & -1 & 1 \end{vmatrix}$

18. $\begin{vmatrix} 1 & -1 & 2 & 0 & 0 \\ 3 & 1 & 4 & 0 & 0 \\ 2 & -1 & 5 & 0 & 0 \\ 0 & 0 & 0 & 2 & 3 \\ 0 & 0 & 0 & -1 & 4 \end{vmatrix}$

19. $\begin{vmatrix} a & 0 & 0 & 0 & 0 \\ 0 & 0 & b & 0 & 0 \\ 0 & 0 & 0 & 0 & c \\ 0 & 0 & 0 & d & 0 \\ 0 & e & 0 & 0 & 0 \end{vmatrix}$

20. $\begin{vmatrix} 2 & 5 & -6 & 8 & 0 \\ 0 & 1 & -7 & 6 & 0 \\ 0 & 0 & 0 & 4 & 0 \\ 0 & 2 & 1 & 5 & 1 \\ 4 & -1 & 5 & 3 & 0 \end{vmatrix}$

En los Problemas del 21 al 27 calcule el determinante suponiendo que

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = 8$$

21. $\begin{vmatrix} a_{31} & a_{32} & a_{33} \\ a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \end{vmatrix}$

22. $\begin{vmatrix} a_{31} & a_{32} & a_{33} \\ a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{vmatrix}$

23. $\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ 2a_{21} & 2a_{22} & 2a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$

24. $\begin{vmatrix} -3a_{11} & -3a_{12} & -3a_{13} \\ 2a_{21} & 2a_{22} & 2a_{23} \\ 5a_{31} & 5a_{32} & 5a_{33} \end{vmatrix}$

25. $\begin{vmatrix} a_{11} & 2a_{13} & a_{12} \\ a_{21} & 2a_{23} & a_{22} \\ a_{31} & 2a_{33} & a_{32} \end{vmatrix}$

26. $\begin{vmatrix} a_{11}-a_{12} & a_{12} & a_{13} \\ a_{21}-a_{22} & a_{22} & a_{23} \\ a_{31}-a_{32} & a_{32} & a_{33} \end{vmatrix}$

27. $\begin{vmatrix} 2a_{11}-3a_{21} & 2a_{12}-3a_{22} & 2a_{13}-3a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{21} & a_{22} & a_{23} \end{vmatrix}$

28. Usando la Propiedad 2, muestre que si α es un número y A es una matriz de $n \times n$, entonces $\det \alpha A = \alpha^n \det A$.

★ 29. Demuestre que

$$\begin{vmatrix} 1+x_1 & x_2 & x_3 & \cdots & x_n \\ x_1 & 1+x_2 & x_3 & \cdots & x_n \\ x_1 & x_2 & 1+x_3 & \cdots & x_n \\ \vdots & \vdots & \vdots & & \vdots \\ x_1 & x_2 & x_3 & \cdots & 1+x_n \end{vmatrix} = 1+x_1+x_2+\cdots+x_n$$

★ 30. Una matriz es *antisimétrica* si $A^t = -A$. Si A es una matriz antisimétrica de $n \times n$, muestre que $\det A = (-1)^n \det A$.

31. Usando el resultado del Problema 30, muestre que si A es una matriz antisimétrica de $n \times n$ y n es impar, entonces $\det A = 0$.

32. Una matriz A se llama *ortogonal* si A es invertible y $A^{-1} = A^t$. Muestre que si A es ortogonal, entonces $\det A = \pm 1$.

★★ 33. Supongamos que Δ denota el triángulo en el plano con vértices en (x_1, y_1), (x_2, y_2) y (x_3, y_3). Demuestre que el área del triángulo está dada por

$$\text{Área de } \Delta = \pm\tfrac{1}{2} \begin{vmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{vmatrix}$$

¿En qué condiciones este determinante es igual a cero?

★★ 34. Tres rectas, las cuales no son paralelas dos a dos, determinan un triángulo en el plano. Suponga que las rectas están dadas por

$$a_{11}x + a_{12}y + a_{13} = 0$$
$$a_{21}x + a_{22}y + a_{23} = 0$$
$$a_{31}x + a_{32}y + a_{33} = 0$$

Muestre que el área determinada por las rectas es

$$\frac{\pm 1}{2A_{13}A_{23}A_{33}} \begin{vmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{vmatrix}$$

35. El *determinante de Vandermonde** de 3×3 está dado por

$$D_3 = \begin{vmatrix} 1 & 1 & 1 \\ a_1 & a_2 & a_3 \\ a_1^2 & a_2^2 & a_3^2 \end{vmatrix}$$

Muestre que $D_3 = (a_2 - a_1)(a_3 - a_1)(a_3 - a_2)$.

36. $D_4 = \begin{vmatrix} 1 & 1 & 1 & 1 \\ a_1 & a_2 & a_3 & a_4 \\ a_1^2 & a_2^2 & a_3^2 & a_4^2 \\ a_1^3 & a_2^3 & a_3^3 & a_4^3 \end{vmatrix}$ es el determinante de Vandermonde de 4×4.

Demuestre que $D_4 = (a_2 - a_1)(a_3 - a_1)(a_4 - a_1)(a_3 - a_2)(a_4 - a_2)(a_4 - a_3)$.

★★ **37.** **a.** Defina el determinante de Vandermonde D_n de $n \times n$.

b. Muestre que $D_n = \prod_{\substack{i=1 \\ j>i}}^{n} (a_j - a_i)$, donde Π representa a la palabra "producto".

Note que el producto del Problema 36 se puede escribir $\prod_{\substack{i=1 \\ j>i}}^{4} (a_j - a_i)$.

38. Sean $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ y $B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$.

a. Escriba el producto AB.
b. Calcule $\det A$, $\det B$ y $\det AB$.
c. Muestre que $\det AB = (\det A)(\det B)$

39. Una matriz A de $n \times n$ se llama *nulipotente* si $A^k = 0$, la matriz cero de $n \times n$ para algún entero $k \geq 1$. Demuestre que las siguientes matrices son nulipotentes y encuentre el entero k más pequeño para el cual $A^k = 0$.

a. $\begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}$ **b.** $\begin{pmatrix} 0 & 1 & 3 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{pmatrix}$

40. Demuestre que si A es nulipotente, entonces $\det A = 0$.

41. La matriz A se dice *idempotente* si $A^2 = A$. ¿Cuáles son los posibles valores de $\det A$ si A es idempotente?

2.3 Si el tiempo lo permite: Demostración de tres teoremas importantes

Teorema 1 **Teorema Básico.** Sea $A = (a_{ij})$ una matriz de $n \times n$. Entonces

$$\det A = a_{11}A_{11} + a_{12}A_{12} + \cdots + a_{1n}A_{1n}$$
$$= a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in} \quad (1)$$
$$= a_{1j}A_{1j} + a_{2j}A_{2j} + \cdots + a_{nj}A_{nj} \quad (2)$$

para $i = 1, 2, \ldots, n$ y $j = 1, 2, \ldots, n$.

* A.T. Vandermonde (1735-1796) fue un matemático francés.

Nota. La primera igualdad es la Definición 2.1.3 del determinante por desarrollo o expansión de cofactores en el primer renglón; la segunda igualdad dice que desarrollando por cofactores en cualquier otro renglón obtenemos el determinante, y la tercera igualdad expresa que desarrollando por cofactores en cualquier columna también se obtiene el determinante.

Demostración Probaremos la igualdad (1) por inducción matemática. Para la matriz de 2×2 $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, primero se desarrolla el primer renglón por cofactores: $\det A = a_{11}A_{11} + a_{12}A_{12} = a_{11}(a_{22}) + a_{12}(-a_{21}) = a_{11}a_{22} - a_{12}a_{21}$. Análogamente, si se hace lo mismo en el segundo renglón, obtenemos $a_{21}A_{21} + a_{22}A_{22} = a_{21}(-a_{12}) + a_{22}(a_{11}) = a_{11}a_{22} - a_{12}a_{21}$. De este modo se produce el mismo resultado desarrollando en cualquier renglón de una matriz de 2×2, lo cual demuestra la igualdad (1) en el caso de 2×2.

Supongamos ahora que la igualdad (1) es válida para todas las matrices de $(n-1) \times (n-1)$. Hay que demostrar que también es válida para matrices de $n \times n$. Nuestro procedimiento será desarrollar por cofactores en el primero y el i-ésimo renglón y mostrar que los desarrollos son idénticos. Si se desarrolla o expande en el primer renglón, entonces un término característico en la expansión del cofactor es

$$a_{1k}A_{1k} = (-1)^{1+k}a_{1k}|M_{1k}| \quad (3)$$

Observemos que éste es el único lugar en la expansión de $|A|$ en donde aparece el término a_{1k}, ya que otro término característico es $a_{1m}A_{1m} = (-1)^{1+m}|M_{1m}|$, $k \neq m$, y M_{1m} la obtenemos eliminando el primer renglón y la m-ésima columna de A (y a_{1k} está en el primer renglón de A). Puesto que M_{1k} es una matriz de $(n-1) \times (n-1)$, podemos, por la hipótesis de inducción, calcular $|M_{1k}|$ expandiendo en el i-ésimo renglón de A (el cual es el $(i-1)$-ésimo renglón de M_{1k}). Un término característico en esta expansión es

$$a_{il} \text{ (cofactor de } a_{il} \text{ en } M_{1k}) \qquad (k \neq l) \quad (4)$$

Por las mismas razones, éste es el único término en la expansión de $|M_{1k}|$, en el i-ésimo renglón de A, en el que aparece a_{il}. Sustituyendo (4) en (3),

$$(-1)^{1+k}a_{1k}a_{il} \text{ (cofactor de } a_{il} \text{ en } M_{1k}) \qquad (k \neq l) \quad (5)$$

es la única vez que aparece el término $a_{1k}a_{il}$, en la expansión por cofactores del det A, en el primer renglón.

Ahora, si expandimos por cofactores en el i-ésimo renglón de A (en donde $i \neq 1$), un término característico es

$$(-1)^{i+1}a_{il}|M_{il}| \quad (6)$$

y un término característico en la expansión de $|M_{il}|$ en el primer renglón es

$$a_{1k} \text{ (cofactor de } a_{1k} \text{ en } M_{il}) \qquad (k \neq l) \quad (7)$$

y sustituyendo (7) en (6), encontramos que la única vez que aparece el término $a_{il}a_{1k}$ en la expansión de $\det A$, en el i-ésimo renglón es

$$(-1)^{i+l}a_{1k}a_{il}\ (\text{cofactor de } a_{1k} \text{ en } M_{il}) \qquad (k \neq l) \tag{8}$$

Si logramos demostrar que las expresiones (5) y (8) son las mismas, entonces (1) quedará demostrado, ya que en (5) es la única vez que aparece $a_{1k}a_{il}$, expandiendo en el primer renglón, y en (8) es la única vez que aparece $a_{1k}a_{il}$ en la expansión del renglón i-ésimo, siendo k, i y l arbitrarios. Esto mostrará que las sumas de los términos en las expansiones del primero y el i-ésimo renglón son las mismas.

Supongamos que $M_{1i,kl}$ denote la matriz de $(n-2)\times(n-2)$ obtenida eliminando el primero y el i-ésimo renglones y la k-ésima y la l-ésima columnas de A. (A esto se le llama *menor de segundo orden* de A.) Supongamos primero que $k<l$. Entonces

$$M_{1k} = \begin{pmatrix} a_{21} & \cdots & a_{2,k-1} & a_{2,k+1} & \cdots & a_{2l} & \cdots & a_{2n} \\ \vdots & & \vdots & \vdots & & \vdots & & \vdots \\ a_{i1} & \cdots & a_{i,k-1} & a_{i,k+1} & \cdots & a_{il} & \cdots & a_{in} \\ \vdots & & \vdots & \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,k-1} & a_{n,k+1} & \cdots & a_{nl} & \cdots & a_{nn} \end{pmatrix} \tag{9}$$

$$M_{il} = \begin{pmatrix} a_{11} & \cdots & a_{1k} & \cdots & a_{1,l-1} & a_{1,l+1} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,k} & \cdots & a_{i-1,l-1} & a_{i-1,l+1} & \cdots & a_{i-1,n} \\ a_{i+1,1} & \cdots & a_{i+1,k} & \cdots & a_{i+1,l-1} & a_{i+1,l+1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nk} & \cdots & a_{n,l-1} & a_{n,l+1} & \cdots & a_{nn} \end{pmatrix} \tag{10}$$

De (9) y (10), vemos que

$$\text{Cofactor de } a_{il} \text{ en } M_{1k} = (-1)^{(i-1)+(l-1)}|M_{1i,kl}| \tag{11}$$

$$\text{Cofactor de } a_{1k} \text{ en } M_{il} = (-1)^{1+k}\,|M_{1i,kl}| \tag{12}$$

Así (5) se convierte en

$$(-1)^{1+k}a_{1k}a_{il}(-1)^{(i-1)+(l-1)}|M_{1i,kl}| = (-1)^{i+k+l-1}a_{1k}a_{il}\,|M_{1i,kl}| \tag{13}$$

y (8) en

$$(-1)^{i+l}a_{1k}a_{il}(-1)^{1+k}|M_{1i,kl}| = (-1)^{i+k+l+1}a_{1k}a_{il}|M_{1i,kl}| \tag{14}$$

Pero $(-1)^{i+k+l-1} = (-1)^{i+k+l+1}$, entonces los segundos miembros de las Ecuaciones (13) y (14) son iguales. De aquí que las expresiones (5) y (8) son iguales y (1) queda demostrado en el caso $k < l$. Si $k > l$, entonces, por un razonamiento análogo,

$$\text{Cofactor de } a_{il} \text{ en } M_{1k} = (-1)^{(i-1)+l}\,|M_{1i,kl}|$$

$$\text{Cofactor de } a_{1k} \text{ en } M_{il} = (-1)^{1+(k-1)}\,|M_{1i,kl}|$$

así que (5) se convierte en

$$(-1)^{1+k}a_{1k}a_{il}(-1)^{(i-1)+l}|M_{1i,kl}| = (-1)^{i+k+l}a_{1k}a_{il}|M_{1i,kl}|$$

y (8) en

$$(-1)^{i+l}a_{1k}a_{il}(-1)^{1+k}|M_{1i,kl}| = (-1)^{i+k+l+1}a_{1k}a_{il}|M_{1i,kl}|$$

Esto completa la demostración de la Ecuación (1).

Para demostrar la Ecuación (2) seguiremos un procedimiento análogo. Si expandimos en la k-ésima y la l-ésima columnas, encontramos que la única vez en que aparece el término $a_{1k}a_{il}$ estará dada por (5) y (8). (Vea Problemas 1 y 2.) Esto muestra que la expansión por cofactores en dos columnas cualesquiera es la misma y que cada una es igual al desarrollo en cualquier renglón. Lo anterior completa la demostración. ■

Ahora deseamos demostrar que para cualesquiera dos matrices de $n \times n$ denominadas A y B, $\det AB = \det A \det B$. La demostración es difícil e implica un cierto número de pasos. Haremos uso de un determinado número de resultados sobre matrices elementales que se demostraron en la Sección 1.10.

Se comenzará calculando los determinantes de las matrices elementales.

Lema 1 Sea E una matriz elemental.

i. Si $E = P_{ij}$, entonces $\det E = -1$ (15)
ii. Si $E = A_{ij}(c)$, entonces $\det E = 1$ (16)
iii. Si $E = M_i(c)$, entonces $\det E = c$ (17)

Demostración **i.** $\det I = 1$. E se obtiene de I intercambiando el renglón i con el j. De la Propiedad 4, $\det E = (-1)\det I = -1$.
ii. E se obtiene de I multiplicando por c el i-ésimo renglón de I y sumándolo al renglón j. Por la Propiedad 7, $\det E = \det I = 1$.
iii. E se obtiene de I multiplicando por c el i-ésimo renglón de I. Por la Propiedad 2, $\det E = c\det I = c$. ■

Lema 2 Sea B una matriz $n \times n$ y sea E una matriz elemental. Entonces

$$\det EB = \det E \det B \tag{18}$$

La demostración del lema se deriva del Lema 1 y de los resultados relativos a las operaciones elementales de renglones discutidas en la Sección 2.2. Los pasos en la demostración se indican en los Problemas del 6 al 8.

Lema 3 Sea T una matriz triangular superior. Entonces T es invertible si y sólo si $\det T \neq 0$.

Demostración Sea

$$T = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{pmatrix} \tag{19}$$

Del Teorema 2.1.1,

$$\det T = a_{11}a_{22}\cdots a_{nn} \tag{20}$$

Así que $\det T \neq 0$ si y sólo si cada una de sus componentes diagonales es distinta de cero.

Si $\det T \neq 0$, entonces T puede ser reducida por renglones del siguiente modo:

i. Para $i = 1, 2, \ldots, n$ divídase entre $a_{ii} \neq 0$ la i-ésima fila de T para obtener

$$\begin{pmatrix} 1 & a'_{12} & \cdots & a'_{1n} \\ 0 & 1 & \cdots & a'_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

ii. Utilice el 1 en la j-ésima componente diagonal para hacer cero cada componente por encima de él en la j-ésima columna.

Así pues, T es equivalente por renglón a I y por lo tanto es invertible según el Teorema 1.8.6 (el Teorema Resumen).

Supóngase que $\det T = 0$. Entonces al menos un elemento diagonal es cero. Sea a_{ii} el primer elemento tal. Entonces T puede ser escrita como

$$T = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1,i-1} & a_{1i} & a_{1,i+1} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2,i-1} & a_{2i} & a_{2,i+1} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{i-1,i-1} & a_{i-1,i} & a_{i-1,i+1} & \cdots & a_{i-1,n} \\ 0 & 0 & \cdots & 0 & 0 & a_{i,i+1} & \cdots & a_{in} \\ 0 & 0 & \cdots & 0 & 0 & a_{i+1,i+1} & \cdots & a_{i+1,n} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & a_{nn} \end{pmatrix} \qquad (21)$$

Considérese el sistema homogéneo

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1,i-1}x_{i-1} + \quad a_{1i}x_i &= 0 \\ a_{22}x_2 + \cdots + a_{2,i-1}x_{i-1} + \quad a_{2i}x_i &= 0 \\ \vdots \qquad\qquad & \vdots \\ a_{i-1,i-1}x_{i-1} + a_{i-1,i}x_i &= 0 \end{aligned}$$

Este es un sistema de $i - 1$ ecuaciones en i incógnitas. Por el Teorema 1.7.1, el sistema tiene soluciones no triviales.

Sea $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_i \end{pmatrix}$ una solución tal que no todas las $x_1, x_2, \ldots, x_i$ sean cero. Sea

$$\mathbf{x}^* = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_i \\ 0 \\ \vdots \\ 0 \end{pmatrix} \left.\begin{matrix} \\ \\ \\ \end{matrix}\right\} n - i \text{ ceros}$$

De (21), y por el modo en que son elegidas las x, vemos que $T\mathbf{x}^* = \mathbf{0}$; esto es, la ecuación $T\mathbf{x} = \mathbf{0}$ tiene una solución no trivial. Utilizando el Teorema 1.8.6 de nuevo (parte *ii*), concluimos que T no es invertible. ■

El siguiente teorema es muy importante.

Teorema 2 Sea A una matriz de $n \times n$. Entonces A es invertible si y sólo si $\det A \neq 0$.

Demostración Del Teorema 1.10.7 sabemos que es posible encontrar matrices $E_1, E_2, \ldots, E_m$ y una matriz triangular T tales que

$$A = E_1E_2 \cdots E_mT \qquad (22)$$

Utilizando m veces el Lema 2, se ve que

$$\begin{aligned} \det A &= \det E_1 \det(E_2E_3 \cdots E_mT) \\ &= \det E_1 \det E_2 \det(E_3 \cdots E_mT) \\ &\ \ \vdots \\ &= \det E_1 \det E_2 \cdots \det E_{m-1} \det(E_mT) \end{aligned}$$

o finalmente,

$$\det A = \det E_1 \det E_2 \cdots \det E_{m-1} \det E_m \det T \qquad (23)$$

Por el Lema 1, $\det E_i \neq 0$. Concluimos que $\det A \neq 0$ si y sólo si $\det T \neq 0$.

Ahora supóngase que A es invertible. Entonces, usando (22) y el hecho de que toda matriz elemental es invertible, podemos escribir T como el producto de matrices invertibles. Así pues, T es invertible, y por el Lema 3, su determinante no se anula. Por tanto, $\det A$ tampoco se anula.

Si $\det A \neq 0$, entonces, por (23), $\det T \neq 0$, y por el Lema 3, T es invertible. El lado derecho de (22) es el producto de matrices invertibles, y por lo tanto, A lo es también. Con esto termina la demostración. ■

Ahora, finalmente, demostraremos el resultado principal.

Teorema 3 Sean A y B matrices $n \times n$. Entonces

$$\det AB = \det A \det B \qquad (24)$$

Demostración *Caso 1:* $\det A = \det B = 0$. Entonces, por el Teorema 2, B no es invertible, y por lo mismo y por el Teorema 1.8.6, existe un n-vector $\mathbf{x} \neq \mathbf{0}$ tal que $B\mathbf{x} = \mathbf{0}$. Entonces, $(AB)\mathbf{x} = A(B\mathbf{x}) = A\mathbf{0} = \mathbf{0}$, lo cual, en virtud del Teorema 1.8.6, vuelve a decir que AB no es invertible. Por el Teorema 2,

$$0 = \det AB = 0 \cdot 0 = \det A \det B.$$

Caso 2: $\det A = 0$ y $\det B \neq 0$; A no es invertible, de modo que existe un n-vector $\mathbf{y} \neq \mathbf{0}$ tal que $A\mathbf{y} = \mathbf{0}$. Como $\det B \neq 0$, B es invertible y existe un vector único $\mathbf{x} \neq \mathbf{0}$ tal que $B\mathbf{x} = \mathbf{y}$. Así pues $AB\mathbf{x} = A(B\mathbf{x}) = A\mathbf{y} = 0$ lo cual muestra que AB no es invertible, y por lo mismo,

$$\det AB = 0 = 0 \det B = \det A \det B.$$

Caso 3: $\det A \neq 0$, A es invertible y puede ser escrita como el producto de matrices elementales:

$$A = E_1E_2 \cdots E_m$$

Entonces,

$$AB = E_1E_2 \cdots E_mB$$

Usando el resultado del Lema 2 en forma repetida, vemos que

$$\begin{aligned}\det AB &= \det(E_1E_2 \cdots E_mB)\\ &= \det E_1 \det E_2 \cdots \det E_m \det B\\ &= \det(E_1E_2 \cdots E_m) \det B\\ &= \det A \det B. \blacksquare\end{aligned}$$

Problemas 2.3

1. Demuestre que si A se expande en su k-ésima columna, entonces la única vez en que aparece el término $a_{1k}a_{il}$ la da la Ecuación (5).
2. Pruebe que si A se expande en su l-ésima columna, entonces la única vez en que aparece el término $a_{1k}a_{il}$ la da la Ecuación (8).
3. Muestre que si A se expande en su k-ésima columna, entonces la única vez en que aparece el término $a_{ik}a_{jl}$ es $(-1)^{i+k}a_{ik}a_{jl}$ (cofactor de a_{jl} en M_{ik}) para $l \neq k$.
4. Sea $A = \begin{pmatrix} 1 & 5 & 7 \\ 2 & -1 & 3 \\ 4 & 5 & -2 \end{pmatrix}$. Calcule $\det A$ expandiendo en cada uno de los renglones y columnas.
5. Haga lo mismo para la matriz $A = \begin{pmatrix} 1 & -1 & 4 \\ 0 & 1 & 5 \\ -3 & 7 & 2 \end{pmatrix}$.
6. Sea $E = P_{ij}$ y sea B una matriz $n \times n$. Demuestre que $\det EB = \det E \det B$. [*Sugerencia:* Describa la matriz EB y luego utilice la Ecuación (15) y la Propiedad 4.]
7. Sea $E = A_{ij}(c)$ y sea B una matriz $n \times n$. Demuestre que $\det EB = \det E \det B$. [*Sugerencia:* Describa la matriz EB y luego utilice la Ecuación (16) y la Propiedad 7.]
8. Sea $E = M_i(c)$ y sea B una matriz $n \times n$. Demuestre que $\det EB = \det E \det B$. [*Sugerencia:* Describa la matriz EB y luego utilice la Ecuación (7) y la Propiedad 2.]

2.4 Determinantes e inversas

En esta sección veremos cómo se pueden calcular las matrices inversas usando determinantes. Más aún, completaremos el trabajo que iniciamos en el Capítulo 1, de demostrar el importante Teorema Resumen 2.7.6 (véanse Teoremas 1.8.6 y 1.10.4), mostrando la equivalencia de varias propiedades de las matrices. Empezaremos con un resultado sencillo.

Teorema 1 Si A es invertible, entonces $\det A \neq 0$ y

$$\det A^{-1} = \frac{1}{\det A} \tag{1}$$

Demostración Del Teorema 2.3.2, $\det A \neq 0$. Del Teorema 2.2.4,

$$1 = \det I = \det AA^{-1} = \det A \det A^{-1} \tag{2}$$

lo cual implica que

$$\det A^{-1} = 1/\det A. \blacksquare$$

Antes de usar determinantes para calcular inversas, necesitamos definir la *adjunta* de una matriz $A = (a_{ij})$. Sea $B = (A_{ij})$, la matriz de cofactores de A. (Recuerde que un cofactor, como se definió en la página 98, es un número.) Entonces

$$B = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix} \tag{3}$$

Definición 1 **Adjunta.** Sea A una matriz de $n \times n$ y sea B, dada por (3), la matriz de sus cofactores. Entonces la *adjunta de A*, escrita como adj A, es la *transpuesta* de la matriz B de $n \times n$. Esto es,

$$\text{adj } A = B^t = \begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix} \tag{4}$$

Ejemplo 1 Sea $A = \begin{pmatrix} 2 & 4 & 3 \\ 0 & 1 & -1 \\ 3 & 5 & 7 \end{pmatrix}$. Calcule adj A.

Solución Tenemos $A_{11} = \begin{vmatrix} 1 & -1 \\ 5 & 7 \end{vmatrix} = 12$, $A_{12} = -\begin{vmatrix} 0 & -1 \\ 3 & 7 \end{vmatrix} = -3$, $A_{13} = -3$, $A_{21} = -13$, $A_{22} = 5$, $A_{23} = 2$, $A_{31} = -7$, $A_{32} = 2$, y $A_{33} = 2$. Así $B = \begin{pmatrix} 12 & -3 & -3 \\ -13 & 5 & 2 \\ -7 & 2 & 2 \end{pmatrix}$ y adj $A = B^t = \begin{pmatrix} 12 & -13 & -7 \\ -3 & 5 & 2 \\ -3 & 2 & 2 \end{pmatrix}$.

Ejemplo 2 Sea

$$A=\begin{pmatrix}1 & -3 & 0 & -2\\ 3 & -12 & -2 & -6\\ -2 & 10 & 2 & 5\\ -1 & 6 & 1 & 3\end{pmatrix}$$

Calcule adj A.

Solución Este caso requiere un poco más de trabajo puesto que tenemos que calcular dieciséis determinantes de 3×3. Por ejemplo tenemos $A_{12}=-\begin{vmatrix}3 & -2 & -6\\ -2 & 2 & 5\\ -1 & 1 & 3\end{vmatrix}=-1$, $A_{24}=\begin{vmatrix}1 & -3 & 0\\ -2 & 10 & 2\\ -1 & 6 & 1\end{vmatrix}=-2$ y $A_{43}=-\begin{vmatrix}1 & -3 & -2\\ 3 & -12 & -6\\ -2 & 10 & 5\end{vmatrix}=3$. Completando estos cálculos, encontramos que

$$B=\begin{pmatrix}0 & -1 & 0 & 2\\ -1 & 1 & -1 & -2\\ 0 & 2 & -3 & -3\\ -2 & -2 & 3 & 2\end{pmatrix}$$

$$\text{adj } A=B^t=\begin{pmatrix}0 & -1 & 0 & -2\\ -1 & 1 & 2 & -2\\ 0 & -1 & -3 & 3\\ 2 & -2 & -3 & 2\end{pmatrix}$$

Ejemplo 3 Sea $A=\begin{pmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{pmatrix}$. Entonces adj $A=\begin{pmatrix}A_{11} & A_{21}\\ A_{12} & A_{22}\end{pmatrix}=\begin{pmatrix}a_{22} & -a_{12}\\ -a_{21} & a_{11}\end{pmatrix}$.

Advertencia. Cuando calcule la adjunta de una matriz, no olvide transponer la matriz de cofactores.

Teorema 2 Sea A una matriz de $n\times n$. Entonces

$$(A)(\text{adj } A)=\begin{pmatrix}\det A & 0 & 0 & \cdots & 0\\ 0 & \det A & 0 & \cdots & 0\\ 0 & 0 & \det A & \cdots & 0\\ \vdots & \vdots & \vdots & & \vdots\\ 0 & 0 & 0 & \cdots & \det A\end{pmatrix}=(\det A)I \qquad (5)$$

Demostración Sea $C=(c_{ij})=(A)(\text{adj } A)$. Entonces

$$C=\begin{pmatrix}a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & & \vdots\\ a_{n1} & a_{n2} & \cdots & a_{nn}\end{pmatrix}\begin{pmatrix}A_{11} & A_{21} & \cdots & A_{n1}\\ A_{12} & A_{22} & \cdots & A_{n2}\\ \vdots & \vdots & & \vdots\\ A_{1n} & A_{2n} & \cdots & A_{nn}\end{pmatrix} \qquad (6)$$

Tenemos

$$c_{ij}=(i\text{-ésimo renglón de } A)\cdot(j\text{-ésima columna de adj } A)$$

$$=(a_{i1}\quad a_{i2}\cdots a_{in})\cdot\begin{pmatrix}A_{j1}\\ A_{j2}\\ \vdots\\ A_{jn}\end{pmatrix}$$

Así

$$c_{ij}=a_{i1}A_{j1}+a_{i2}A_{j2}+\cdots+a_{in}A_{jn} \qquad (7)$$

Ahora, si $i=j$, la suma en (7) es igual a $a_{i1}A_{i1}+a_{i2}A_{i2}+\cdots+a_{in}A_{in}$, lo cual es la expansión de $\det A$ en el i-ésimo renglón de A. Por otro lado, si $i\neq j$, entonces el Teorema 2.2.2 la suma en (7) es igual a cero. Así

$$c_{ij}=\begin{cases}\det A & \text{si } i=j\\ 0 & \text{si } i\neq j\end{cases}$$

Esto demuestra el teorema. ■

Podemos ahora establecer el resultado principal.

Teorema 3 Sea A una matriz de $n\times n$. Entonces A es invertible si y sólo si $\det A\neq 0$. Si $\det A\neq 0$, entonces

$$A^{-1}=\frac{1}{\det A}\text{adj } A \qquad (8)$$

Note que el Teorema 1.8.4 para matrices de 2×2 es un caso especial de este teorema.

Demostración La primera parte del teorema es el Teorema 2.3.2. Si $\det A\neq 0$, entonces,

$$(A)\left(\frac{1}{\det A}\text{adj } A\right)=\frac{1}{\det A}[A(\text{adj } A)]\overset{\text{Teorema 2}}{=}\frac{1}{\det A}(\det A)I=I$$

Pero, por el Teorema 1.8.8, si $AB=I$, entonces $B=A^{-1}$. De modo que $(1/\det A)\text{ adj } A=A^{-1}$. ■

Ejemplo 4 Sea $A=\begin{pmatrix}2 & 4 & 3\\ 0 & 1 & -1\\ 3 & 5 & 7\end{pmatrix}$. Determine si A es invertible y, si lo es, calcule A^{-1}.

Solución Puesto que $\det A = 3 \neq 0$, vemos que A es invertible. Del Ejemplo 1,

$$\text{adj } A = \begin{pmatrix} 12 & -13 & -7 \\ -3 & 5 & 2 \\ -3 & 2 & 2 \end{pmatrix}.$$

Así

$$A^{-1} = \tfrac{1}{3}\begin{pmatrix} 12 & -13 & -7 \\ -3 & 5 & 2 \\ -3 & 2 & 2 \end{pmatrix} = \begin{pmatrix} 4 & -\frac{13}{3} & -\frac{7}{3} \\ -1 & \frac{5}{3} & \frac{2}{3} \\ -1 & \frac{2}{3} & \frac{2}{3} \end{pmatrix}$$

Verificación $$A^{-1}A = \tfrac{1}{3}\begin{pmatrix} 12 & -13 & -7 \\ -3 & 5 & 2 \\ -3 & 2 & 2 \end{pmatrix}\begin{pmatrix} 2 & 4 & 3 \\ 0 & 1 & -1 \\ 3 & 5 & 7 \end{pmatrix} = \tfrac{1}{3}\begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix} = I$$

Ejemplo 5 Sea

$$A = \begin{pmatrix} 1 & -3 & 0 & -2 \\ 3 & -12 & -2 & -6 \\ -2 & 10 & 2 & 5 \\ -1 & 6 & 1 & 3 \end{pmatrix}$$

Determine si A es invertible y, en tal caso, calcule A^{-1}.

Solución Usando las propiedades de los determinantes, calculamos

$$\begin{vmatrix} 1 & -3 & 0 & -2 \\ 3 & -12 & -2 & -6 \\ -2 & 10 & 2 & 5 \\ -1 & 6 & 1 & 3 \end{vmatrix}$$

Multiplique la primera columna por 3 y por 2 y súmela a la segunda y cuarta columnas, respectivamente.

$$= \begin{vmatrix} 1 & 0 & 0 & 0 \\ 3 & -3 & -2 & 0 \\ -2 & 4 & 2 & 1 \\ -1 & 3 & 1 & 1 \end{vmatrix}$$

Expanda en el primer renglón.

$$= \begin{vmatrix} -3 & -2 & 0 \\ 4 & 2 & 1 \\ 3 & 1 & 1 \end{vmatrix} = -1$$

Así, $\det A = -1 \neq 0$ y A^{-1} existe. Por el Ejemplo 2, tenemos

$$\text{adj } A = \begin{pmatrix} 0 & -1 & 0 & -2 \\ -1 & 1 & 2 & -2 \\ 0 & -1 & -3 & 3 \\ 2 & -2 & -3 & 3 \end{pmatrix}$$

Así

$$A^{-1} = \frac{1}{-1}\begin{pmatrix} 0 & -1 & 0 & -2 \\ -1 & 1 & 2 & -2 \\ 0 & -1 & -3 & 3 \\ 2 & -2 & -3 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 2 \\ 1 & -1 & -2 & 2 \\ 0 & 1 & 3 & -3 \\ -2 & 2 & 3 & -2 \end{pmatrix}$$

Verificación $$AA^{-1} = \begin{pmatrix} 1 & -3 & 0 & -2 \\ 3 & -12 & -2 & -6 \\ -2 & 10 & 2 & 5 \\ -1 & 6 & 1 & 3 \end{pmatrix}\begin{pmatrix} 0 & 1 & 0 & 2 \\ 1 & -1 & -2 & 2 \\ 0 & 1 & 3 & -3 \\ -2 & 2 & 3 & -2 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Nota. Como se habrá observado, si $n > 3$ generalmente es más fácil evaluar A^{-1} reduciendo por renglón y después usando adj A puesto que, aun para el caso de 4×4, es necesario calcular 17 determinantes (16 para la adjunta y además $\det A$). Sin embargo, el Teorema 3 es muy importante ya que, antes de que se haga cualquier reducción por renglón, el cálculo de $\det A$ (en caso de que pueda hacerse fácilmente) nos dirá si A^{-1} existe o no.

Habíamos visto el Teorema Resumen (Teoremas 1.2.1, 1.8.6 y 1.10.4) en la Sección 1.10. Este es el teorema que relaciona muchos de los conceptos expuestos en los primeros dos capítulos de este libro.

Teorema 4 ***Teorema resumen, Versión 4.*** Sea A una matriz de $n \times n$. Entonces, cada uno de los siguientes seis enunciados implica los otros cinco. Es decir, que si uno es verdadero, todos los demás lo son.

i. A es invertible.
ii. La única solución al sistema homogéneo $A\mathbf{x} = \mathbf{0}$ es la solución trivial ($\mathbf{x} = \mathbf{0}$).
iii. El sistema $A\mathbf{x} = \mathbf{b}$ tiene una solución única para cada n-vector $\mathbf{b}$.
iv. A es equivalente por renglones a la matriz identidad I_n de $n \times n$.
v. A es el producto de matrices elementales.
vi. $\det A \neq 0$.

Demostración En el Teorema 1.8.6 demostramos la equivalencia de las partes (*i*), (*ii*), (*iii*) y (*iv*). En el Teorema 1.10.3 vimos la equivalencia de (*i*) y (*v*). El Teorema 1 (o el 2.3.2) prueba la equivalencia de (*i*) y (*vi*). ■

Problemas 2.4

En los Problemas del 1 *al* 12 *use los métodos de esta sección para determinar si la matriz dada es invertible. Si es así, calcule la inversa.*

1. $\begin{pmatrix} 3 & 2 \\ 1 & 2 \end{pmatrix}$ 2. $\begin{pmatrix} 3 & 6 \\ -4 & -8 \end{pmatrix}$ 3. $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

4. $\begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 3 \\ 5 & 5 & 1 \end{pmatrix}$ 5. $\begin{pmatrix} 3 & 2 & 1 \\ 0 & 2 & 2 \\ 0 & 1 & -1 \end{pmatrix}$ 6. $\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$.

7. $\begin{pmatrix} 1 & 2 & 3 \\ 1 & 1 & 2 \\ 0 & 1 & 2 \end{pmatrix}$ 8. $\begin{pmatrix} 3 & 1 & 0 \\ 1 & -1 & 2 \\ 1 & 1 & 1 \end{pmatrix}$ 9. $\begin{pmatrix} 2 & -1 & 4 \\ -1 & 0 & 5 \\ 19 & -7 & 3 \end{pmatrix}$

10. $\begin{pmatrix} 1 & 6 & 2 \\ -2 & 3 & 5 \\ 7 & 12 & -4 \end{pmatrix}$ 11. $\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & -1 & 2 \\ 1 & -1 & 2 & 1 \\ 1 & 3 & 3 & 2 \end{pmatrix}$ 12. $\begin{pmatrix} 1 & -3 & 0 & -2 \\ 3 & -12 & -2 & -6 \\ -2 & 10 & 2 & 5 \\ -1 & 6 & 1 & 3 \end{pmatrix}$

13. Muestre que una matriz A de $n \times n$ es invertible si y sólo si A^t es invertible.

14. Para $A = \begin{pmatrix} 1 & 1 \\ 2 & 5 \end{pmatrix}$, verifique que $\det A^{-1} = 1/\det A$.

15. Para $A = \begin{pmatrix} 1 & -1 & 3 \\ 4 & 1 & 6 \\ 2 & 0 & -2 \end{pmatrix}$, verifique que $\det A^{-1} = 1/\det A$.

16. ¿Para qué valores de α la matriz $\begin{pmatrix} \alpha & -3 \\ 4 & 1-\alpha \end{pmatrix}$ no es invertible?

17. ¿Para qué valores de α la matriz $\begin{pmatrix} -\alpha & \alpha-1 & \alpha+1 \\ 1 & 2 & 3 \\ 2-\alpha & \alpha+3 & \alpha+7 \end{pmatrix}$ no tiene una inversa?

18. Suponga que la matriz A de $n \times n$ no es invertible. Muestre que $(A)(\text{adj } A)$ es la matriz cero.

2.5 Regla de Cramer

En esta sección examinaremos un método antiguo para resolver sistemas con el mismo número de incógnitas y de ecuaciones. Considere el sistema de n ecuaciones en n incógnitas.

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \vdots \qquad \vdots \qquad \quad \vdots \quad &\;\; \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n \end{aligned} \tag{1}$$

el cual puede ser escrito en la forma

$$A\mathbf{x} = \mathbf{b} \tag{2}$$

Supongamos que $\det A \neq 0$. Entonces el sistema (2) tiene una única solución dada por $\mathbf{x} = A^{-1}\mathbf{b}$. Podemos desarrollar un método para encontrar esa solución sin reducción por renglones y sin calcular A^{-1}.

Sea $D = \det A$. Definimos n matrices nuevas:

$$A_1 = \begin{pmatrix} b_1 & a_{12} & \cdots & a_{1n} \\ b_2 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ b_n & a_{n2} & \cdots & a_{nn} \end{pmatrix}, \quad A_2 = \begin{pmatrix} a_{11} & b_1 & \cdots & a_{1n} \\ a_{21} & b_2 & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & b_n & \cdots & a_{nn} \end{pmatrix}, \ldots,$$

$$A_n = \begin{pmatrix} a_{11} & a_{12} & \cdots & b_1 \\ a_{21} & a_{22} & \cdots & b_2 \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & b_n \end{pmatrix}$$

Es decir, A_i es la matriz que se obtiene sustituyendo la i-ésima columna de A por $\mathbf{b}$. Finalmente, sea $D_1 = \det A_1, D_2 = \det A_2, \ldots, D_n = \det A_n$.

Teorema 4 ***Regla de Cramer.**** Sea A una matriz de $n \times n$ y suponga que $\det A \neq 0$. Entonces, la solución única para el sistema $A\mathbf{x} = \mathbf{b}$ está dada por

$$x_1 = \frac{D_1}{D},\ x_2 = \frac{D_2}{D}, \ldots, x_i = \frac{D_i}{D}, \ldots, x_n = \frac{D_n}{D} \tag{3}$$

Demostración La solución de $A\mathbf{x} = \mathbf{b}$ es $\mathbf{x} = A^{-1}\mathbf{b}$. Pero

$$A^{-1}\mathbf{b} = \frac{1}{D}(\text{adj } A)\mathbf{b} = \frac{1}{D}\begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix}\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \tag{4}$$

Ahora $(\text{adj } A)\mathbf{b}$ es un n-vector cuya j-ésima componente es

$$(A_{1j}\ A_{2j} \cdots A_{nj}) \cdot \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = b_1A_{1j} + b_2A_{2j} + \cdots + b_nA_{nj} \tag{5}$$

---

* Llamada así por el matemático suizo Gabriel Cramer (1704-1752). Cramer publicó la regla en 1750 en su *Introduction to the Analysis of Lines of Algebraic Curves*. En realidad hay evidencia de que la regla era conocida desde 1729 por Colin Maclaurin (1698-1746), quien fue, probablemente, el matemático británico más destacado en los años posteriores a la muerte de Newton. La regla de Cramer es uno de los resultados más famosos en la historia de las Matemáticas. Durante casi doscientos años, fue esencial en la enseñanza del Álgebra y de la teoría de ecuaciones. Debido al enorme número de cálculos que implica su uso, la regla es poco utilizada hoy día. Sin embargo, el resultado fue muy importante en su momento.

Consideremos la matriz A_j:

$$A_j = \begin{pmatrix} a_{11} & a_{12} & \cdots & b_1 & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & b_2 & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & b_n & \cdots & a_{nn} \end{pmatrix} \tag{6}$$

j-ésima columna

Si expandimos el determinante de A_j en su j-ésima columna,

$$D_j = b_1 \text{ (cofactor de } b_1) + b_2 \text{ (cofactor de } b_2) + \cdots + b_n \text{ (cofactor de } b_n) \tag{7}$$

Pero para encontrar el cofactor de b_i, por ejemplo, eliminamos el i-ésimo renglón y la j-ésima columna de A_j (puesto que b_i se encuentra en la j-ésima columna de A_j). Pero la j-ésima columna de A_j es $\mathbf{b}$ y con ésta eliminada, tenemos simplemente el ij-ésimo menor M_{ij} de A. Así

$$\text{Cofactor de } b_i \text{ en } A_j = A_{ij}$$

de modo que (7) se convierte en

$$D_j = b_1A_{1j} + b_2A_{2j} + \cdots + b_nA_{nj} \tag{8}$$

Pero esto es igual que el lado derecho de (5). Así, el i-ésimo componente de (adj A)$\mathbf{b}$ es D_i y

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = A^{-1}\mathbf{b} = \frac{1}{D}(\text{adj } A)\mathbf{b} = \frac{1}{D}\begin{pmatrix} D_1 \\ D_2 \\ \vdots \\ D_n \end{pmatrix} = \begin{pmatrix} D_1/D \\ D_2/D \\ \vdots \\ D_n/D \end{pmatrix}$$

La demostración está completa. ■

Ejemplo 1 Usando la regla de Cramer resuelva el sistema

$$\begin{aligned} 2x_1 + 4x_2 + 6x_3 &= 18 \\ 4x_1 + 5x_2 + 6x_3 &= 24 \\ 3x_1 + x_2 - 2x_3 &= 4 \end{aligned} \tag{9}$$

Solución Ya hemos resuelto este sistema usando reducción por renglones en el Ejemplo 1.6.1. Podríamos resolverlo también calculando A^{-1} (Ejemplo 1.8.6) y después encontrando $A^{-1}\mathbf{b}$. Ahora lo resolveremos usando la regla de Cramer. Primero

$$D = \begin{vmatrix} 2 & 4 & 6 \\ 4 & 5 & 6 \\ 3 & 1 & -2 \end{vmatrix} = 6 \neq 0$$

de manera que el sistema (9) tiene una única solución. Entonces

$$D_1 = \begin{vmatrix} 18 & 4 & 6 \\ 24 & 5 & 6 \\ 4 & 1 & -2 \end{vmatrix} = 24, \quad D_2 = \begin{vmatrix} 2 & 18 & 6 \\ 4 & 24 & 6 \\ 3 & 4 & -2 \end{vmatrix} = -12 \quad \text{y}$$

$$D_3 = \begin{vmatrix} 2 & 4 & 18 \\ 4 & 5 & 24 \\ 3 & 1 & 4 \end{vmatrix} = 18. \quad \text{. Por lo tanto } x_1 = \frac{D_1}{D} = \frac{24}{6} = 4,$$

$$x_2 = \frac{D_2}{D} = -\frac{12}{6} = -2 \quad \text{y} \quad x_3 = \frac{D_3}{D} = \frac{18}{6} = 3.$$

Ejemplo 2 Muestre que el sistema

$$\begin{aligned} x_1 + 3x_2 + 5x_3 + 2x_4 &= 2 \\ -x_2 + 3x_3 + 4x_4 &= 0 \\ 2x_1 + x_2 + 9x_3 + 6x_4 &= -3 \\ 3x_1 + 2x_2 + 4x_3 + 8x_4 &= -1 \end{aligned} \tag{10}$$

tiene solución única y encuéntrela usando la regla de Cramer.

Solución Vimos en el Ejemplo 2.2.10 que

$$|A| = \begin{vmatrix} 1 & 3 & 5 & 2 \\ 0 & -1 & 3 & 4 \\ 2 & 1 & 9 & 6 \\ 3 & 2 & 4 & 8 \end{vmatrix} = 160 \neq 0$$

De modo que el sistema tiene una única solución. Para encontrarla calculamos: $D_1 = -464$; $D_2 = 280$; $D_3 = -56$; $D_4 = 112$. Así $x_1 = D_1/D = -464/160$, $x_2 = D_2/D = 280/160$, $x_3 = D_3/D = -56/160$ y $x_4 = D_4/D = 112/160$. Estas soluciones pueden ser verificadas por sustitución directa en el sistema (10).

Problemas 2.5

En los Problemas del 1 *al* 9 *resuelva el sistema dado, usando la regla de Cramer.*

1. $2x_1 + 3x_2 = -1$
   $-7x_1 + 4x_2 = 47$
2. $3x_1 - x_2 = 0$
   $4x_1 + 2x_2 = 5$
3. $2x_1 + x_2 + x_3 = 6$
   $3x_1 - 2x_2 - 3x_3 = 5$
   $8x_1 + 2x_2 + 5x_3 = 11$
4. $x_1 + x_2 + x_3 = 8$
   $4x_2 - x_3 = -2$
   $3x_1 - x_2 + 2x_3 = 0$
5. $2x_1 + 2x_2 + x_3 = 7$
   $x_1 + 2x_2 - x_3 = 0$
   $-x_1 + x_2 + 3x_3 = 1$
6. $2x_1 + 5x_2 - x_3 = -1$
   $4x_1 + x_2 + 3x_3 = 3$
   $-2x_1 + 2x_2 = 0$

7. $\begin{aligned} 2x_1 + x_2 - x_3 &= 4 \\ x_1 + x_3 &= 2 \\ -x_2 + 5x_3 &= 1 \end{aligned}$

8. $\begin{aligned} x_1 + x_2 + x_3 + x_4 &= 6 \\ 2x_1 - x_3 - x_4 &= 4 \\ 3x_3 + 6x_4 &= 3 \\ x_1 - x_4 &= 5 \end{aligned}$

9. $\begin{aligned} x_1 - x_4 &= 7 \\ 2x_2 + x_3 &= 2 \\ 4x_1 - x_2 &= -3 \\ 3x_3 - 5x_4 &= 2 \end{aligned}$

★10. Considere el triángulo en la Figura 2.2.

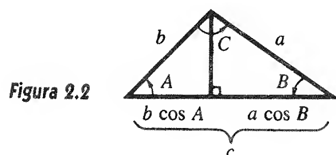


Figura 2.2

a. Usando trigonometría elemental, muestre que

$$\begin{aligned} c \cos A + a \cos C &= b \\ b \cos A + a \cos B &= c \\ c \cos B + b \cos C &= a \end{aligned}$$

b. Si el sistema de la parte (*a*) se considera como un sistema de tres ecuaciones en las tres incógnitas cos *A*, cos *B* y cos *C*, pruebe que el determinante del sistema no es cero.

c. Use la regla de Cramer para resolver el sistema para cos *C*.

d. Aplique (*c*) para demostrar la *ley de los cosenos*: $c^2 = a^2 + b^2 - 2ab \cos C$.

Ejercicios de repaso • Capítulo 2

En los Ejercicios del 1 *al* 8 *calcule el determinante.*

1. $\begin{vmatrix} -1 & 2 \\ 0 & 4 \end{vmatrix}$

2. $\begin{vmatrix} -3 & 5 \\ -7 & 4 \end{vmatrix}$

3. $\begin{vmatrix} 1 & -2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{vmatrix}$

4. $\begin{vmatrix} 5 & 0 & 0 \\ 6 & 2 & 0 \\ 10 & 100 & 6 \end{vmatrix}$

5. $\begin{vmatrix} 1 & -1 & 2 \\ 3 & 4 & 2 \\ -2 & 3 & 4 \end{vmatrix}$

6. $\begin{vmatrix} 3 & 1 & -2 \\ 4 & 0 & 5 \\ -6 & 1 & 3 \end{vmatrix}$

7. $\begin{vmatrix} 1 & -1 & 2 & 3 \\ 4 & 0 & 2 & 5 \\ -1 & 2 & 3 & 7 \\ 5 & 1 & 0 & 4 \end{vmatrix}$

8. $\begin{vmatrix} 3 & 15 & 17 & 19 \\ 0 & 2 & 21 & 60 \\ 0 & 0 & 1 & 50 \\ 0 & 0 & 0 & -1 \end{vmatrix}$

En los Ejercicios del 9 *al* 14 *use determinantes para calcular la inversa, si la hay.*

9. $\begin{pmatrix} -3 & 4 \\ 2 & 1 \end{pmatrix}$

10. $\begin{pmatrix} 3 & -5 & 7 \\ 0 & 2 & 4 \\ 0 & 0 & -3 \end{pmatrix}$

11. $\begin{pmatrix} 1 & -1 & 2 \\ 3 & 1 & 4 \\ 5 & -1 & 8 \end{pmatrix}$

12. $\begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$

13. $\begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & -1 & 3 & 0 \\ 1 & 0 & 0 & -2 \\ 3 & 0 & -1 & 0 \end{pmatrix}$

14. $\begin{pmatrix} 3 & -1 & 2 & 4 \\ 1 & 1 & 0 & 3 \\ -2 & 4 & 1 & 5 \\ 6 & -4 & 1 & 2 \end{pmatrix}$

En los Ejercicios del 15 *al* 18 *resuelva el sistema empleando la regla de Cramer.*

15. $\begin{aligned} 2x_1 - x_2 &= 3 \\ 3x_1 + 2x_2 &= 5 \end{aligned}$

16. $\begin{aligned} x_1 - x_2 + x_3 &= 7 \\ 2x_1 - 5x_3 &= 4 \\ 3x_2 - x_3 &= 2 \end{aligned}$

17. $\begin{aligned} 2x_1 + 3x_2 - x_3 &= 5 \\ -x_1 + 2x_2 + 3x_3 &= 0 \\ 4x_1 - x_2 + x_3 &= -1 \end{aligned}$

18. $\begin{aligned} x_1 - x_3 + x_4 &= 7 \\ 2x_2 + 2x_3 - 3x_4 &= -1 \\ 4x_1 - x_2 - x_3 &= 0 \\ -2x_1 + x_2 + 4x_3 &= 2 \end{aligned}$

CAPÍTULO **3**

Vectores en $\mathbb{R}^2$ y $\mathbb{R}^3$

En la Sección 1.3 se definieron los vectores columna y renglón como conjuntos ordenados de n números reales. En el siguiente capítulo definiremos otros tipos de conjuntos de vectores, llamados *espacios vectoriales*.

El estudio de espacios vectoriales arbitrarios requiere, inicialmente, de un esfuerzo considerable de abstracción. Por esa misma razón es útil poseer un "almacén" de vectores fácilmente visualizables, los cuales pueden ser utilizados en los ejemplos.

En este capítulo discutiremos las propiedades básicas de los vectores en el plano xy y en el espacio tridimensional real. Los estudiantes que han cursado cálculo de varias variables ya habrán visto este material con anterioridad. Si éste es el caso, este capítulo debe ser expuesto en forma breve y como un repaso. Para otros casos, la cobertura de este material proveerá al estudiante de ejemplos que hacen mucho más comprensible el material de los Capítulos 4 y 5.

3.1 Vectores en el plano

Como lo definimos en la Sección 1.3, $\mathbb{R}^2$ es el conjunto de vectores (x_1, x_2), con x_1 y x_2 números reales. Como cada punto del plano se puede escribir en la forma (x, y), es evidente que cualquier punto del plano se puede ver como un vector en $\mathbb{R}^2$ y viceversa. Por tanto usaremos indistintamente los términos "el plano" y "$\mathbb{R}^2$". Sin embargo, para varias de las aplicaciones físicas (incluyendo los conceptos de fuerza, velocidad, aceleración y cantidad de movimiento) es importante pensar en un vector no como un punto, sino como un objeto que tiene "magnitud" y "dirección". Veremos ahora cómo se hace esto.

Sean P y Q dos puntos en el plano. Entonces el *segmento de recta dirigido* de P a Q, denotado por $\overrightarrow{PQ}$, es el segmento rectilíneo que va de P a Q (vea Figura 3.1*a*). Notemos que los segmentos de recta dirigidos $\overrightarrow{PQ}$ y $\overrightarrow{QP}$ son diferentes pues apuntan en direcciones opuestas (Figura 3.1*b*).

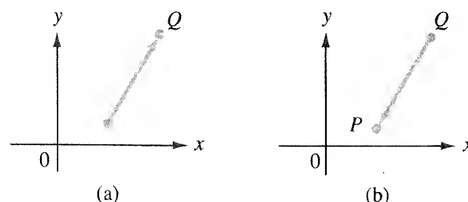


Figura 3.1

El punto P en el segmento de recta dirigido $\overrightarrow{PQ}$ se conoce como *punto inicial* del segmento, y el punto Q como *punto terminal*. Las dos propiedades principales de un segmento de recta dirigido son su magnitud (longitud) y su dirección. Si dos segmentos dirigidos $\overrightarrow{PQ}$ y $\overrightarrow{RS}$ tienen igual magnitud y dirección decimos que son *equivalentes*, sin que interese su ubicación con respecto al origen. Todos los segmentos dirigidos de la Figura 3.2 son equivalentes.

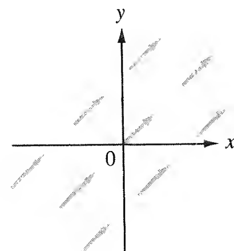


Figura 3.2

Definición 1 ***Definición geométrica de un vector.*** El conjunto de todos los segmentos de recta dirigidos equivalentes a un segmento dirigido dado, se llama *vector*. Cualquier segmento de recta dirigido en ese conjunto se conoce como un *representante* del vector.

Observación. Todos los segmentos de recta dirigidos en la Figura 3.2 son representantes del mismo vector.

De la Definición 1 vemos que un vector dado **v** se puede representar de diferentes maneras. Sea $\overrightarrow{PQ}$ un representante de **v**. Entonces, sin cambiar su magnitud ni su dirección podemos mover $\overrightarrow{PQ}$ paralelamente hasta que su punto inicial quede en el origen. Hemos obtenido así el segmento de recta dirigido $\overrightarrow{OR}$ que es otro representante del vector **v** (vea Figura 3.3). Ahora supongamos que R tiene coordenadas cartesianas (a, b). Entonces podemos describir el segmento de recta dirigido $\overrightarrow{OR}$, por las coordenadas (a, b). Esto es, $\overrightarrow{OR}$ es el segmento dirigido con punto inicial $(0, 0)$ y punto terminal (a, b). Como un representante de un vector sirve tan bien como otro podemos expresar el vector **v** como (a, b).

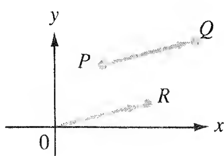


Figura 3.3

Definición 2 ***Definición algebraica de un vector.*** Un *vector* **v** en el plano xy es un par ordenado de números reales (a, b). Los números a y b se conocen como las *componentes* del vector **v**. El *vector cero* es $(0, 0)$.

Observación 1. Con esta definición, un punto en el plano xy puede considerarse como un vector que se inicia en el origen y termina en ese punto.

Observación 2. El vector cero tiene magnitud cero. Por tanto, como el punto inicial y el terminal coinciden decimos que el vector cero *no tiene dirección*.

Observación 3. Enfatizamos que las Definiciones 1 y 2 describen exactamente los mismos objetos. Cada punto de vista (geométrico y algebraico) tiene sus ventajas. La Definición 2 es la definición de un vector con dos componentes que hemos venido usando hasta ahora.

Como un vector es realmente un conjunto de segmentos de recta equivalentes, definimos la *magnitud* o *longitud* de un vector como la magnitud de cualquiera de sus representantes, y su *dirección*, como la de cualquiera de sus representantes. Si usamos el representante $\overline{OR}$ y consideramos el vector $\mathbf{v} = (a, b)$, se tiene que

$$|\mathbf{v}| = \text{magnitud de } \mathbf{v} = \sqrt{a^2+b^2} \qquad (1)$$

Esto se sigue del teorema de Pitágoras (vea Figura 3.4). Hemos usado la notación $|\mathbf{v}|$ para simbolizar la magnitud de **v**. Notemos que $|\mathbf{v}|$ es un *escalar*.

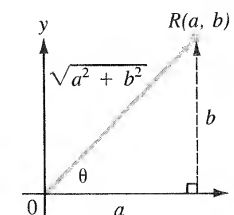


Figura 3.4

Ejemplo 1 Calcule las magnitudes de los vectores (i) $(2, 2)$; (ii) $(2, 2\sqrt{3})$; (iii) $(-2\sqrt{3}, 2)$; (iv) $(-3, -3)$; (v) $(6, -6)$.

Solución

i. $|\mathbf{v}| = \sqrt{2^2+2^2} = \sqrt{8} = 2\sqrt{2}$
ii. $|\mathbf{v}| = \sqrt{2^2+(2\sqrt{3})^2} = 4$
iii. $|\mathbf{v}| = \sqrt{(-2\sqrt{3})^2+2^2} = 4$
iv. $|\mathbf{v}| = \sqrt{(-3)^2+(-3)^2} = \sqrt{18} = 3\sqrt{2}$
v. $|\mathbf{v}| = \sqrt{6^2+(-6)^2} = \sqrt{72} = 6\sqrt{2}$

Dirección de un vector Ahora definimos la *dirección* del vector $\mathbf{v} = (a, b)$ como el ángulo θ (medido en radianes) que forma el vector con la parte positiva del eje x. Por convención escogemos θ tal que $0 \le \theta < 2\pi$. De la Figura 3.4 se observa que si $a \ne 0$, entonces

$$\tan \theta = \frac{b}{a} \tag{2}$$

Ejemplo 2 Calcule las direcciones de los vectores del Ejemplo 1.

Solución Estos cinco vectores están representados en la Figura 3.5.

i. Aquí $\mathbf{v}$ está en el primer cuadrante y como $\tan \theta = 2/2 = 1$, $\theta = \pi/4$.
ii. Aquí $\theta = \tan^{-1} 2\sqrt{3}/2 = \tan^{-1}\sqrt{3} = \pi/3$ (pues $\mathbf{v}$ está en el primer cuadrante).
iii. Vemos que $\mathbf{v}$ está en el segundo cuadrante y, como $\tan^{-1}2/2\sqrt{3} = \tan^{-1}1/\sqrt{3} = \pi/6$, resulta de la Figura 3.5c que $\theta = \pi - (\pi/6) = 5\pi/6$.
iv. Aquí $\mathbf{v}$ está en el tercer cuadrante y, como $\tan^{-1} 1 = \pi/4$, tenemos que $\theta = \pi + (\pi/4) = 5\pi/4$.
v. Como $\mathbf{v}$ está en el cuarto cuadrante y $\tan^{-1}(-1) = -\pi/4$, se tiene que $\theta = 2\pi - (\pi/4) = 7\pi/4$.

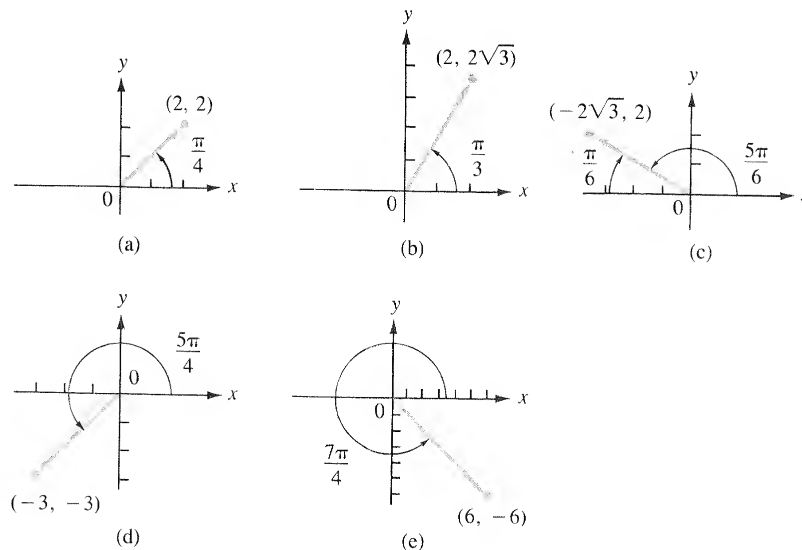


Figura 3.5

En la Sección 1.3 definimos la suma de vectores y la multiplicación por un escalar. ¿Qué significan geométricamente estos conceptos? Empezamos con

la multiplicación por un escalar. Si $\mathbf{v} = (a, b)$ entonces $\alpha\mathbf{v} = (\alpha a, \alpha b)$. De manera que

$$|\alpha \mathbf{v}| = \sqrt{\alpha^2 a^2 + \alpha^2 b^2} = |\alpha|\sqrt{a^2 + b^2} = |\alpha|\,|\mathbf{v}| \tag{3}$$

Esto es:

Multiplicar un vector por un escalar tiene el efecto de multiplicar la magnitud del vector por el valor absoluto de ese escalar.

Más aún, si $\alpha > 0$, entonces $\alpha\mathbf{v}$ está en el mismo cuadrante que $\mathbf{v}$ y, por tanto, la dirección de $\alpha\mathbf{v}$ es la *misma* que la dirección de $\mathbf{v}$ pues $\tan^{-1}(\alpha b/\alpha a) = \tan^{-1}(b/a)$. Si $\alpha < 0$, entonces $\alpha\mathbf{v}$ apunta en la dirección opuesta a la de $\mathbf{v}$. En otras palabras:

$$\begin{aligned} &\text{Dirección de } \alpha\mathbf{v} = \text{dirección de } \mathbf{v}, \text{ si } \alpha > 0 \\ &\text{Dirección de } \alpha\mathbf{v} = \text{dirección de } \mathbf{v} + \pi, \text{ si } \alpha < 0 \end{aligned} \tag{4}$$

Ejemplo 3 Sea $\mathbf{v} = (1, 1)$. Entonces $|\mathbf{v}| = \sqrt{1+1} = \sqrt{2}$ y $|2\mathbf{v}| = |(2, 2)| = \sqrt{2^2 + 2^2} = \sqrt{8} = 2\sqrt{2} = 2|\mathbf{v}|$. Más aún, $|-2\mathbf{v}| = \sqrt{(-2)^2 + (-2)^2} = 2\sqrt{2} = 2|\mathbf{v}|$. Además, la dirección de $2\mathbf{v}$ es $\pi/4$ mientras que la dirección de $-2\mathbf{v}$ es $5\pi/4$ (vea Figura 3.6).

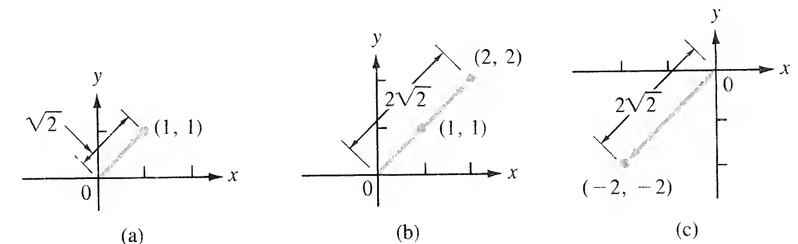


Figura 3.6

Supongamos ahora que se suman los vectores $\mathbf{u} = (a_1, b_1)$ y $\mathbf{v} = (a_2, b_2)$ como en la Figura 3.7. En el croquis vemos que el vector $\mathbf{u} + \mathbf{v} = (a_1 + a_2, b_1 + b_2)$ se puede obtener trasladando el representante del vector

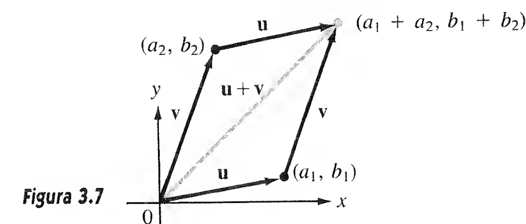


Figura 3.7

v de manera que su punto inicial coincida con el punto terminal (a_1, b_1) del vector **u**. Podemos obtener así el vector **u** + **v** dibujando un paralelogramo con uno de sus vértices en el origen y lados **u** y **v**. Entonces **u** + **v** es el vector que va desde el origen a lo largo de la diagonal del paralelogramo.

Nota. Como la recta es la mínima distancia entre dos puntos, se deduce inmediatamente de la Figura 3.7 que

$$|\mathbf{u}+\mathbf{v}| \leq |\mathbf{u}|+|\mathbf{v}| \qquad (5)$$

Por razones obvias la desigualdad (5) se conoce como la *desigualdad del triángulo.*

Podemos usar la Figura 3.7 también para obtener una representación geométrica del vector **u** − **v**. Como **u** = **u** − **v** + **v**, el vector **u** − **v** es el vector que debe ser sumado a **v** para obtener **u**. Este hecho se ilustra en la Figura 3.8*a*. Un hecho similar se ilustra en la Figura 3.8*b*.

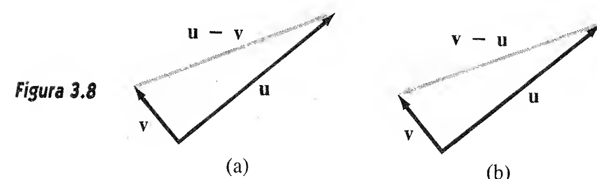


Figura 3.8

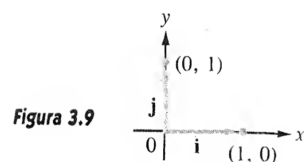


Figura 3.9

Existen dos vectores especiales en $\mathbb{R}^2$ que nos permiten representar otros vectores de $\mathbb{R}^2$ en una forma conveniente. Denotaremos el vector (1, 0) con el símbolo **i** y el vector (0, 1) con el símbolo **j** (vea Figura 3.9).* Si **v** = (a, b) es otro vector del plano entonces, como $(a, b) = a(1, 0) + b(0, 1)$,

$$\mathbf{v}=(a, b)=a\mathbf{i}+b\mathbf{j} \qquad (6)$$

Con esta representación decimos que **v** está *resuelto en sus componentes vertical y horizontal.* Los vectores **i** y **j** tienen dos propiedades:

* *Nota histórica:* Los símbolos **i** y **j** fueron usados por primera vez por Hamilton, quien definió el cuaternión como una cantidad de la forma $a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$, donde a es la "parte escalar" y $b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$ la "parte vectorial". En la Sección 3.3 se expresarán los vectores en el espacio según la forma $b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$.

i. Ninguno de ellos es múltiplo del otro. (En la terminología del Capítulo 4 se dice que son *linealmente independientes.*)
ii. Cualquier vector **v** se puede escribir en términos de **i** y **j** como en la Ecuación (6).*

En estas dos condiciones se dice que **i** y **j** forman una *base* de $\mathbb{R}^2$. Discutiremos las bases de espacios vectoriales arbitrarios en el Capítulo 4.

Definiremos ahora una clase de vector que es muy útil en ciertas aplicaciones.

Definición 3 ***Vector unitario.*** Un *vector unitario* **u** es un vector de magnitud igual a 1.

Ejemplo 4 El vector $\mathbf{u}=(1/2)\mathbf{i}+(\sqrt{3}/2)\mathbf{j}$ es un vector unitario pues

$$|\mathbf{u}|=\sqrt{\left(\frac{1}{2}\right)^2+\left(\frac{\sqrt{3}}{2}\right)^2}=\sqrt{\frac{1}{4}+\frac{3}{4}}=1$$

Sea $\mathbf{u}=a\mathbf{i}+b\mathbf{j}$ un vector unitario. Entonces $|\mathbf{u}|=\sqrt{a^2+b^2}=1$, de donde $a^2+b^2=1$ y **u** se puede representar por un punto en el círculo unitario (Figura 3.10). Si θ es la dirección de **u**, entonces vemos inmediatamente que $a=\cos\theta$ y $b=\operatorname{sen}\theta$. Así, cualquier vector unitario **u** se puede escribir en la forma

$$\mathbf{u}=(\cos\theta)\mathbf{i}+(\operatorname{sen}\theta)\mathbf{j} \qquad (7)$$

donde θ es la dirección de **u**.

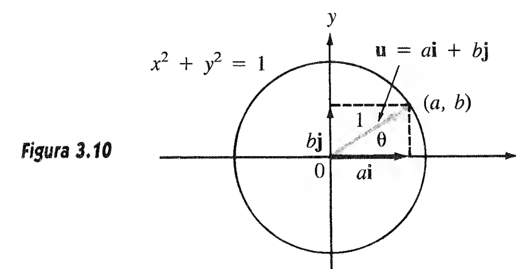


Figura 3.10

Ejemplo 5 El vector unitario $\mathbf{u}=(1/2)\mathbf{i}+(\sqrt{3}/2)\mathbf{j}$ del Ejemplo 4 se puede escribir en la forma (7) con $\theta=\cos^{-1}(1/2)=\pi/3$.

* En la Ecuación (6), decimos que **v** se escribe como una *combinación lineal* de **i** y **j**. Discutiremos el concepto de combinación lineal en la Sección 4.5.

También tenemos:

> Sea $\mathbf{v}$ un vector distinto de cero. Entonces $\mathbf{u} = \mathbf{v}/|\mathbf{v}|$ es un vector unitario con la misma dirección de $\mathbf{v}$ (vea Problema 17).

Ejemplo 6 Encuentre un vector unitario con la misma dirección que $\mathbf{v} = 2\mathbf{i} - 3\mathbf{j}$.

Solución Aquí $|\mathbf{v}| = \sqrt{4+9} = \sqrt{13}$, y así, $\mathbf{u} = \mathbf{v}/|\mathbf{v}| = (2/\sqrt{13})\mathbf{i} - (3/\sqrt{13})\mathbf{j}$ es el vector unitario pedido.

Concluimos esta sección con un resumen de las propiedades de los vectores (Tabla 3.1).

Tabla 3.1

Objeto	Definición intuitiva	Expresión en términos de componentes si $\mathbf{u} = u_1\mathbf{i} + u_2\mathbf{j}$, $\mathbf{v} = v_1\mathbf{i} + v_2\mathbf{j}$, y $\mathbf{u} = (u_1, u_2)$, $\mathbf{v} = (v_1, v_2)$
vector $\mathbf{v}$	Objeto con magnitud y dirección	$v_1\mathbf{i} + v_2\mathbf{j}$ o bien (v_1, v_2)
$\|\mathbf{v}\|$	magnitud de $\mathbf{v}$	$\sqrt{v_1^2 + v_2^2}$
$\alpha\mathbf{v}$	$\mathbf{v}$ $\alpha\mathbf{v}$ (En este esquema $\alpha = 2$)	$\alpha v_1\mathbf{i} + \alpha v_2\mathbf{j}$ o bien $(\alpha v_1, \alpha v_2)$
$-\mathbf{v}$	$\mathbf{v}$ $-\mathbf{v}$	$-v_1\mathbf{i} - v_2\mathbf{j}$ o bien $(-v_1 - v_2)$ o bien $-(v_1, v_2)$
$\mathbf{u}+\mathbf{v}$	$\mathbf{u}+\mathbf{v}$ $\mathbf{v}$ $\mathbf{u}$	$(u_1 + v_1)\mathbf{i} + (u_2 + v_2)\mathbf{j}$ o bien $(u_1 + v_1, u_2 + v_2)$
$\mathbf{u}-\mathbf{v}$	$\mathbf{v}$ $\mathbf{u}-\mathbf{v}$ $\mathbf{u}$	$(u_1 - v_1)\mathbf{i} + (u_2 - v_2)\mathbf{j}$ o bien $(u_1 - v_1, u_2 - v_2)$

Problemas 3.1

En los Problemas del 1 *al* 12 *encuentre la magnitud y dirección del vector dado.*

1. $\mathbf{v} = (4, 4)$
2. $\mathbf{v} = (-4, 4)$
3. $\mathbf{v} = (4, -4)$
4. $\mathbf{v} = (-4, -4)$
5. $\mathbf{v} = (\sqrt{3}, 1)$
6. $\mathbf{v} = (1, \sqrt{3})$
7. $\mathbf{v} = (-1, \sqrt{3})$
8. $\mathbf{v} = (1, -\sqrt{3})$
9. $\mathbf{v} = (-1, -\sqrt{3})$
10. $\mathbf{v} = (1, 2)$
11. $\mathbf{v} = (-5, 8)$
12. $\mathbf{v} = (11, -14)$
13. Sean $\mathbf{u} = (2, 3)$ y $\mathbf{v} = (-5, 4)$. Encuentre: (a) $3\mathbf{u}$; (b) $\mathbf{u} + \mathbf{v}$; (c) $\mathbf{v} - \mathbf{u}$; (d) $2\mathbf{u} - 7\mathbf{v}$. Dibuje estos vectores.

14. Sean $\mathbf{u} = 2\mathbf{i} - 3\mathbf{j}$ y $\mathbf{v} = -4\mathbf{i} + 6\mathbf{j}$. Encuentre: (a) $\mathbf{u} + \mathbf{v}$; (b) $\mathbf{u} - \mathbf{v}$; (c) $3\mathbf{u}$; (d) $-7\mathbf{v}$; (e) $8\mathbf{u} - 3\mathbf{v}$; (f) $4\mathbf{v} - 6\mathbf{u}$. Dibuje estos vectores.
15. Muestre que los vectores $\mathbf{i}$ y $\mathbf{j}$ son vectores unitarios.
16. Muestre que el vector $(1/\sqrt{2})\mathbf{i} + (1/\sqrt{2})\mathbf{j}$ es un vector unitario.
17. Muestre que si $\mathbf{v} = a\mathbf{i} + b\mathbf{j}$, entonces $\mathbf{u} = (a/\sqrt{a^2+b^2})\mathbf{i} + (b/\sqrt{a^2+b^2})\mathbf{j}$ es un vector unitario con la misma dirección de $\mathbf{v}$.

En los Problemas del 18 *al* 21 *encuentre un vector unitario que tenga la misma dirección que el vector dado.*

18. $\mathbf{v} = 2\mathbf{i} + 3\mathbf{j}$
19. $\mathbf{v} = \mathbf{i} - \mathbf{j}$
20. $\mathbf{v} = -3\mathbf{i} + 4\mathbf{j}$
21. $\mathbf{v} = a\mathbf{i} + a\mathbf{j}$; $a \neq 0$.
22. Si $\mathbf{v} = a\mathbf{i} + b\mathbf{j}$, muestre que $a/\sqrt{a^2+b^2} = \cos\theta$ y $b/\sqrt{a^2+b^2} = \operatorname{sen}\theta$, en donde θ es la dirección de $\mathbf{v}$.
23. Si $\mathbf{v} = 2\mathbf{i} - 3\mathbf{j}$, encuentre $\operatorname{sen}\theta$ y $\cos\theta$.
24. Si $\mathbf{v} = -3\mathbf{i} + 8\mathbf{j}$, encuentre $\operatorname{sen}\theta$ y $\cos\theta$.

Un vector $\mathbf{v}$ *tiene la dirección opuesta al vector* $\mathbf{u}$ *si la dirección de* $\mathbf{v}$ *es la dirección de* $\mathbf{u} + \pi$. *En los Problemas del* 25 *al* 28 *encuentre un vector unitario* $\mathbf{v}$ *que tenga la dirección opuesta a la dirección del vector dado* $\mathbf{u}$.

25. $\mathbf{u} = \mathbf{i} + \mathbf{j}$
26. $\mathbf{u} = 2\mathbf{i} - 3\mathbf{j}$
27. $\mathbf{u} = -3\mathbf{i} + 4\mathbf{j}$
28. $\mathbf{u} = -2\mathbf{i} + 3\mathbf{j}$
29. Sean $\mathbf{u} = 2\mathbf{i} - 3\mathbf{j}$ y $\mathbf{v} = -\mathbf{i} + 2\mathbf{j}$. Encuentre un vector unitario en la misma dirección que (a) $\mathbf{u} + \mathbf{v}$; (b) $2\mathbf{u} - 3\mathbf{v}$; (c) $3\mathbf{u} + 8\mathbf{v}$.
30. Sean $P = (c, d)$ y $Q = (c + a, d + b)$. Muestre que la magnitud de $\overrightarrow{PQ}$ es $\sqrt{a^2+b^2}$.
31. Muestre que la dirección de $\overrightarrow{PQ}$ del Problema 30 es la misma que la del vector (a, b). [*Sugerencia:* Si $R = (a, b)$, muestre que la recta que pasa por los puntos P y Q es paralela a la que pasa por los puntos O y R.]

En los Problemas del 32 *al* 35 *encuentre un vector* $\mathbf{v}$ *que tenga la magnitud y dirección dadas.*

32. $|\mathbf{v}| = 3$; $\theta = \pi/6$
33. $|\mathbf{v}| = 8$; $\theta = \pi/3$
34. $|\mathbf{v}| = 1$; $\theta = \pi/4$
35. $|\mathbf{v}| = 6$; $\theta = 2\pi/3$.
36. ★ Muestre algebraicamente (esto es, estrictamente a partir de las definiciones de suma de vectores y de magnitud) que para cualquier par de vectores $\mathbf{u}$ y $\mathbf{v}$, $|\mathbf{u} + \mathbf{v}| \leq |\mathbf{u}| + |\mathbf{v}|$.
37. Muestre que si $\mathbf{u}$ y $\mathbf{v}$ son distintos del vector cero, entonces $|\mathbf{u} + \mathbf{v}| = |\mathbf{u}| + |\mathbf{v}|$ si y sólo si $\mathbf{u}$ es un múltiplo escalar positivo de $\mathbf{v}$.

3.2 El producto escalar y proyecciones en $\mathbb{R}^2$

En la Sección 1.5 definimos el producto escalar de dos vectores. Si $\mathbf{u} = (a_1, b_1)$ y $\mathbf{v} = (a_2, b_2)$, entonces

$$\mathbf{u} \cdot \mathbf{v} = a_1a_2 + b_1b_2 \tag{1}$$

Veremos ahora cómo puede interpretarse geométricamente el producto escalar.

Definición 1 ***Ángulo entre dos vectores.*** Sean **u** y **v** dos vectores distintos de cero. El *ángulo* φ *entre* **u** *y* **v** se define como el menor ángulo* positivo entre los representantes de **u** y **v** que tienen al origen como sus puntos iniciales. Si $\mathbf{u} = \alpha\mathbf{v}$ para algún escalar α, definimos $\varphi = 0$ si $\alpha > 0$ y $\varphi = \pi$ si $\alpha < 0$.

Esta definición se ilustra en la Figura 3.11. Notemos que φ siempre se puede escoger como un ángulo positivo en el intervalo $[0, \pi]$.

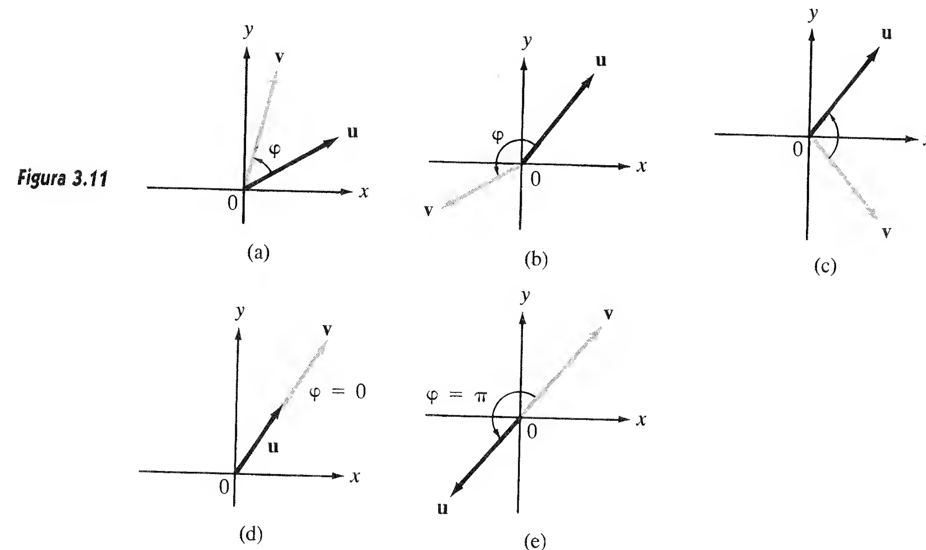


Figura 3.11

Teorema 1 Sean **u** y **v** dos vectores distintos de cero. Si φ es el ángulo entre ellos, entonces

$$\cos\varphi = \frac{\mathbf{u}\cdot\mathbf{v}}{|\mathbf{u}|\,|\mathbf{v}|} \tag{2}$$

Demostración La ley de los cosenos (Problema 2.5.10) establece que en el triángulo de la Figura 3.12.

$$c^2 = a^2 + b^2 - 2ab\cos C \tag{3}$$

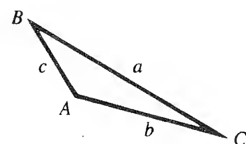


Figura 3.12

Ahora llevamos los representantes de **u** y **v** al origen de forma que

* Este ángulo estará en el intervalo $[0, \pi]$.

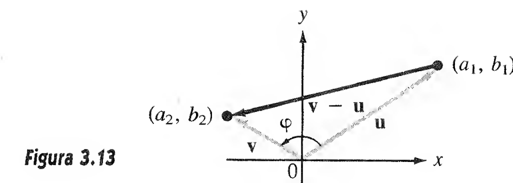


Figura 3.13

$\mathbf{u} = (a_1, b_1)$ y $\mathbf{v} = (a_2, b_2)$ (vea Figura 3.13). Entonces, de la citada ley de los cosenos, $|\mathbf{v}-\mathbf{u}|^2 = |\mathbf{v}|^2 + |\mathbf{u}|^2 - 2|\mathbf{u}|\,|\mathbf{v}|\cos\varphi$. Pero

$$\begin{aligned}|\mathbf{v}-\mathbf{u}|^2 &= (\mathbf{v}-\mathbf{u})\cdot(\mathbf{v}-\mathbf{u}) = \mathbf{v}\cdot\mathbf{v} - 2\mathbf{u}\cdot\mathbf{v} + \mathbf{u}\cdot\mathbf{u}\\ &= |\mathbf{v}|^2 - 2\mathbf{u}\cdot\mathbf{v} + |\mathbf{u}|^2\end{aligned}$$

Así, después de simplificar obtenemos que $-2\mathbf{u}\cdot\mathbf{v} = -2|\mathbf{u}|\,|\mathbf{v}|\cos\varphi$, de donde se deduce el teorema. ■

Observación. Si usamos el Teorema 1 podríamos definir el producto escalar $\mathbf{u}\cdot\mathbf{v}$ por

$$\mathbf{u}\cdot\mathbf{v} = |\mathbf{u}|\,|\mathbf{v}|\cos\varphi$$

Ejemplo 1 Encuentre el coseno del ángulo entre los vectores $\mathbf{u} = 2\mathbf{i} + 3\mathbf{j}$ y $\mathbf{v} = -7\mathbf{i} + \mathbf{j}$.

Solución $\mathbf{u}\cdot\mathbf{v} = -14 + 3 = -11$, $|\mathbf{u}| = \sqrt{2^2+3^2} = \sqrt{13}$, y $|\mathbf{v}| = \sqrt{(-7)^2+1^2} = \sqrt{50}$. Así,

$$\cos\varphi = \frac{\mathbf{u}\cdot\mathbf{v}}{|\mathbf{u}|\,|\mathbf{v}|} = \frac{-11}{\sqrt{13}\sqrt{50}} = \frac{-11}{\sqrt{650}} \approx -0.4315^*$$

Definición 1 ***Vectores paralelos.*** Dos vectores **u** y **v** distintos de cero son *paralelos* si el ángulo entre ellos es cero o bien π.

Ejemplo 2 Muestre que los vectores $\mathbf{u} = (2, -3)$ y $\mathbf{v} = (-4, 6)$ son paralelos.

Solución

$$\cos\varphi = \frac{\mathbf{u}\cdot\mathbf{v}}{|\mathbf{u}|\,|\mathbf{v}|} = \frac{-8-18}{\sqrt{13}\sqrt{52}} = \frac{-26}{\sqrt{13}(2\sqrt{13})} = \frac{-26}{2(13)} = -1$$

Por lo tanto $\varphi = \pi$.

---

* Este número, como otros en el texto, se obtuvo con una calculadora de bolsillo.

Teorema 2 Si $\mathbf{u} \neq \mathbf{0}$, entonces $\mathbf{v} = \alpha\mathbf{u}$ para alguna constante α distinta de cero si y sólo si **u** y **v** son paralelos.

Demostración Esta demostración queda como ejercicio (Problema 35).

Definición 3 ***Vectores ortogonales.*** Se dice que los vectores **u** y **v** distintos de cero son *ortogonales* (o *perpendiculares*) si el ángulo entre ellos es $\pi/2$.

Ejemplo 3 Muestre que los vectores $\mathbf{u} = 3\mathbf{i} - 4\mathbf{j}$ y $\mathbf{v} = 4\mathbf{i} + 3\mathbf{j}$ son ortogonales.

Solución $\mathbf{u} \cdot \mathbf{v} = 3 \cdot 4 - 4 \cdot 3 = 0$. Esto implica que $\cos\varphi = (\mathbf{u} \cdot \mathbf{v})/(|\mathbf{u}|\,|\mathbf{v}|) = 0$. Como φ está en el intervalo $[0, \pi]$, $\varphi = \pi/2$.

Teorema 3 Los vectores **u** y **v** distintos de cero son ortogonales si y sólo si $\mathbf{u} \cdot \mathbf{v} = 0$.

Demostración Esta demostración también queda como ejercicio (Problema 36).

Ejemplo 4 Sea $\mathbf{u} = \mathbf{i} + 4\mathbf{j}$ y $\mathbf{v} = 3\mathbf{i} + \alpha\mathbf{j}$. Determinar α tal que **(i)** **u** y **v** resulten ortogonales; **(ii)** **u** y **v** resulten paralelos.

Solución

i. Tenemos que $\mathbf{u} \cdot \mathbf{v} = 3 + 4\alpha$. Para que **u** y **v** resulten ortogonales, debemos exigir que $\mathbf{u} \cdot \mathbf{v} = 0$. Esto implica que $3 + 4\alpha = 0$, o lo que es lo mismo, $\alpha = -\frac{3}{4}$.

ii. En este caso debemos tener $\varphi = 0$, o bien π, así que $\cos\varphi = \pm 1$. Entonces,

$$\cos\varphi = \frac{\mathbf{u} \cdot \mathbf{v}}{|\mathbf{u}|\,|\mathbf{v}|} = \frac{3+4\alpha}{\sqrt{17}\sqrt{9+\alpha^2}} = \pm 1$$

Elevando al cuadrado ambos lados de esta ecuación, obtenemos $9 + 24\alpha + 16\alpha^2 = 17(9 + \alpha^2) = 153 + 17\alpha^2$. Esto conduce a la ecuación cuadrática $\alpha^2 - 24\alpha + 144 = 0 = (\alpha - 12)^2$, con la solución única $\alpha = 12$.

Un buen número de problemas interesantes implica el concepto de proyección de un vector sobre otro. Antes de dar la definición de esto, demostraremos el siguiente teorema.

Teorema 4 Sea **v** un vector distinto de cero. Entonces para cualquier otro vector **u** el vector $\mathbf{w} = \mathbf{u} - [(\mathbf{u} \cdot \mathbf{v})\mathbf{v}/|\mathbf{v}|^2]$ es ortogonal a **v**.

Demostración

$$\mathbf{w} \cdot \mathbf{v} = \left[\mathbf{u} - \frac{(\mathbf{u} \cdot \mathbf{v})\mathbf{v}}{|\mathbf{v}|^2}\right] \cdot \mathbf{v} = \mathbf{u} \cdot \mathbf{v} - \frac{(\mathbf{u} \cdot \mathbf{v})(\mathbf{v} \cdot \mathbf{v})}{|\mathbf{v}|^2}$$

$$= \mathbf{u} \cdot \mathbf{v} - \frac{(\mathbf{u} \cdot \mathbf{v})|\mathbf{v}|^2}{|\mathbf{v}|^2} = \mathbf{u} \cdot \mathbf{v} - \mathbf{u} \cdot \mathbf{v} = 0 \quad \blacksquare$$

Los vectores **u**, **v** y **w** se ilustran en la Figura 3.14.

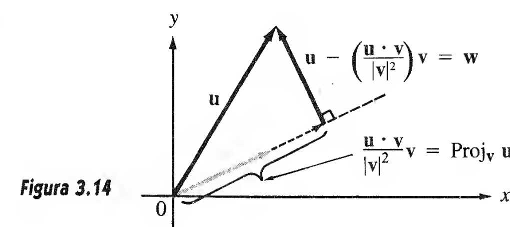


Figura 3.14

Definición 4 ***Proyección.*** Sean **u** y **v** vectores distintos de cero. Entonces la *proyección de* **u** *sobre* **v** es un vector denotado $\text{proy}_{\mathbf{v}}\mathbf{u}$, que se define por

$$\text{proy}_{\mathbf{v}}\,\mathbf{u} = \frac{\mathbf{u} \cdot \mathbf{v}}{|\mathbf{v}|^2}\,\mathbf{v} \tag{4}$$

La *componente* de **u** en la dirección **v** es $\dfrac{\mathbf{u} \cdot \mathbf{v}}{|\mathbf{v}|}$. (5)

Notemos que $\mathbf{v}/|\mathbf{v}|$ es un vector unitario en la dirección de **v**.

Observación 1. De la Figura 3.14 y el hecho de que $\cos\varphi = (\mathbf{u} \cdot \mathbf{v})/(|\mathbf{u}|\,|\mathbf{v}|)$, encontramos que:

> **v** y $\text{proy}_{\mathbf{v}}\mathbf{u}$ tienen (*i*) la misma dirección si $\mathbf{u} \cdot \mathbf{v} > 0$ y (*ii*) direcciones opuestas si $\mathbf{u} \cdot \mathbf{v} < 0$ (vea Figura 3.15).

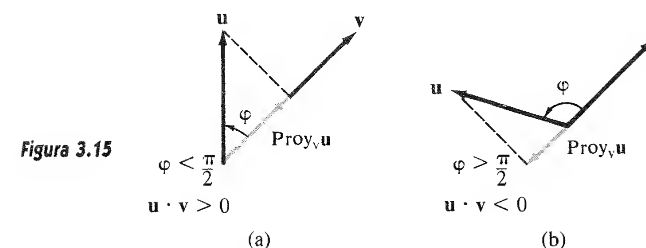


Figura 3.15

Observación 2. Es claro que $\text{proy}_{\mathbf{v}}\mathbf{u}$ se puede considerar la "componente según **v**" del vector **u**.

Observación 3. Si tenemos que **u** y **v** son ortogonales, entonces $\mathbf{u} \cdot \mathbf{v} = 0$ de donde $\text{proy}_\mathbf{v}\mathbf{u} = \mathbf{0}$.

Observación 4. Otra manera de definir proyección es: Si **u** y **v** son vectores distintos de cero, entonces $\text{proy}_\mathbf{v}\mathbf{u}$ es un vector único con las siguientes propiedades:

> i. $\text{proy}_\mathbf{v}\mathbf{u}$ es paralelo a **v**.
> ii. $\mathbf{u} - \text{proy}_\mathbf{v}\mathbf{u}$ es ortogonal a **v**.

Ejemplo 5 Sean $\mathbf{u} = 2\mathbf{i} + 3\mathbf{j}$ y $\mathbf{v} = \mathbf{i} + \mathbf{j}$. Calcule $\text{proy}_\mathbf{v}\mathbf{u}$.

Solución $\text{Proy}_\mathbf{v}\ \mathbf{u} = (\mathbf{u} \cdot \mathbf{v})\mathbf{v}/|\mathbf{v}|^2 = [5/(\sqrt{2})^2]\mathbf{v} = (5/2)\mathbf{i} + (5/2)\mathbf{j}$. (Figura 3.16).

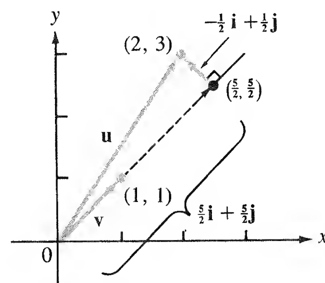


Figura 3.16

Ejemplo 6 Sean $\mathbf{u} = 2\mathbf{i} - 3\mathbf{j}$ y $\mathbf{v} = \mathbf{i} + \mathbf{j}$. Calcule $\text{proy}_\mathbf{v}\mathbf{u}$.

Solución Aquí $(\mathbf{u} \cdot \mathbf{v})/|\mathbf{v}|^2 = -\frac{1}{2}$; por lo tanto, $\text{proy}_\mathbf{v}\ \mathbf{u} = -\frac{1}{2}\mathbf{i} - \frac{1}{2}\mathbf{j}$. (Figura 3.17).

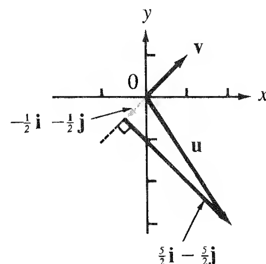


Figura 3.17

Problemas 3.2

En los Problemas del 1 al 8 calcule el producto escalar de los dos vectores y el coseno del ángulo entre ellos.

1. $\mathbf{u} = \mathbf{i} + \mathbf{j}$; $\mathbf{v} = \mathbf{i} - \mathbf{j}$
2. $\mathbf{u} = 3\mathbf{i}$; $\mathbf{v} = -7\mathbf{j}$

3. $\mathbf{u} = -5\mathbf{i}$; $\mathbf{v} = 18\mathbf{j}$
4. $\mathbf{u} = \alpha\mathbf{i}$; $\mathbf{v} = \beta\mathbf{j}$; α, β reales
5. $\mathbf{u} = 2\mathbf{i} + 5\mathbf{j}$; $\mathbf{v} = 5\mathbf{i} + 2\mathbf{j}$
6. $\mathbf{u} = 2\mathbf{i} + 5\mathbf{j}$; $\mathbf{v} = 5\mathbf{i} - 2\mathbf{j}$
7. $\mathbf{u} = -3\mathbf{i} + 4\mathbf{j}$; $\mathbf{v} = -2\mathbf{i} - 7\mathbf{j}$
8. $\mathbf{u} = 4\mathbf{i} + 5\mathbf{j}$; $\mathbf{v} = 5\mathbf{i} - 4\mathbf{j}$
9. Muestre que para cualquier par de números reales α y β, los vectores $\mathbf{u} = \alpha\mathbf{i} + \beta\mathbf{j}$ y $\mathbf{v} = \beta\mathbf{i} - \alpha\mathbf{j}$ son ortogonales.
10. Sean **u**, **v** y **w** tres vectores arbitrarios. Explique por qué el producto $\mathbf{u} \cdot \mathbf{v} \cdot \mathbf{w}$ *no está definido*.

En los Problemas del 11 *al* 16 *diga si los vectores dados son ortogonales, paralelos o ninguna de las dos cosas. Dibuje cada par.*

11. $\mathbf{u} = 3\mathbf{i} + 5\mathbf{j}$; $\mathbf{v} = -6\mathbf{i} - 10\mathbf{j}$
12. $\mathbf{u} = 2\mathbf{i} + 3\mathbf{j}$; $\mathbf{v} = 6\mathbf{i} - 4\mathbf{j}$
13. $\mathbf{u} = 2\mathbf{i} + 3\mathbf{j}$; $\mathbf{v} = 6\mathbf{i} + 4\mathbf{j}$
14. $\mathbf{u} = 2\mathbf{i} + 3\mathbf{j}$; $\mathbf{v} = -6\mathbf{i} + 4\mathbf{j}$
15. $\mathbf{u} = 7\mathbf{i}$; $\mathbf{v} = -23\mathbf{j}$
16. $\mathbf{u} = 2\mathbf{i} - 6\mathbf{j}$; $\mathbf{v} = -\mathbf{i} + 3\mathbf{j}$
17. Sean $\mathbf{u} = 3\mathbf{i} + 4\mathbf{j}$ y $\mathbf{v} = \mathbf{i} + \alpha\mathbf{j}$. Encuentre α tal que:
    a. **u** y **v** sean ortogonales.
    b. **u** y **v** sean paralelos.
    c. El ángulo entre **u** y **v** sea $\pi/4$.
    d. El ángulo entre **u** y **v** sea $\pi/3$.
18. Sea $\mathbf{u} = -2\mathbf{i} + 5\mathbf{j}$ y $\mathbf{v} = \alpha\mathbf{i} - 2\mathbf{j}$. Encuentre α tal que:
    a. **u** y **v** sean ortogonales.
    b. **u** y **v** sean paralelos.
    c. El ángulo entre **u** y **v** sea $2\pi/3$.
    d. El ángulo entre **u** y **v** sea $\pi/3$.
19. Con los datos del Problema 17, muestre que no existe ningún valor de α para el cual **u** y **v** tengan direcciones opuestas.
20. Con los datos del Problema 18 muestre que no existe ningún valor de α para el cual **u** y **v** tengan la misma dirección.

En los Problemas del 21 *al* 30 *calcule* $\text{proy}_\mathbf{v}\mathbf{u}$.

21. $\mathbf{u} = 3\mathbf{i}$; $\mathbf{v} = \mathbf{i} + \mathbf{j}$
22. $\mathbf{u} = -5\mathbf{j}$; $\mathbf{v} = \mathbf{i} + \mathbf{j}$
23. $\mathbf{u} = 2\mathbf{i} + \mathbf{j}$; $\mathbf{v} = \mathbf{i} - 2\mathbf{j}$
24. $\mathbf{u} = 2\mathbf{i} + 3\mathbf{j}$; $\mathbf{v} = 4\mathbf{i} + \mathbf{j}$
25. $\mathbf{u} = \mathbf{i} + \mathbf{j}$; $\mathbf{v} = 2\mathbf{i} - 3\mathbf{j}$
26. $\mathbf{u} = \mathbf{i} + \mathbf{j}$; $\mathbf{v} = 2\mathbf{i} + 3\mathbf{j}$
27. $\mathbf{u} = \alpha\mathbf{i} + \beta\mathbf{j}$; $\mathbf{v} = \mathbf{i} + \mathbf{j}$; α, β reales y positivos
28. $\mathbf{u} = \mathbf{i} + \mathbf{j}$; $\mathbf{v} = \alpha\mathbf{i} + \beta\mathbf{j}$, α, β reales y positivos
29. $\mathbf{u} = \alpha\mathbf{i} - \beta\mathbf{j}$; $\mathbf{v} = \mathbf{i} + \mathbf{j}$; α, β reales y positivos, con $\alpha > \beta$.
30. $\mathbf{u} = \alpha\mathbf{i} - \beta\mathbf{j}$; $\mathbf{v} = \mathbf{i} + \mathbf{j}$; α, β reales y positivos, con $\alpha < \beta$.
31. Sean $\mathbf{u} = a_1\mathbf{i} + b_1\mathbf{j}$ y $\mathbf{v} = a_2\mathbf{i} + b_2\mathbf{j}$. Halle condiciones en a_1, b_1, a_2 y b_2 que aseguren que **v** y $\text{proy}_\mathbf{v}\mathbf{u}$ tengan la misma dirección.
32. En el Problema 31 encuentre una condición que asegure que **v** y $\text{proy}_\mathbf{v}\mathbf{u}$ tengan direcciones opuestas.
33. Sean $P = (2, 3)$, $Q = (5, 7)$, $R = (2, -3)$ y $S = (1, 2)$. Calcule $\text{proy}_{\overrightarrow{PQ}}\overrightarrow{RS}$ y $\text{proy}_{\overrightarrow{RS}}\overrightarrow{PQ}$.
34. Sean $P = (-1, 3)$, $Q = (2, 4)$, $R = (-6, -2)$ y $S = (3, 0)$. Calcule $\text{proy}_{\overrightarrow{PQ}}\overrightarrow{RS}$ y $\text{proy}_{\overrightarrow{RS}}\overrightarrow{PQ}$.
35. Demuestre que dos vectores **u** y **v** distintos de cero son paralelos si y sólo si $\mathbf{v} = \alpha\mathbf{u}$ para alguna constante α. [*Sugerencia:* Muestre que $\cos\varphi = \pm 1$ si y sólo si $\mathbf{v} = \alpha\mathbf{u}$.]
36. Demuestre que **u** y **v** son ortogonales si y sólo si $\mathbf{u} \cdot \mathbf{v} = 0$.
37. Muestre que el vector $\mathbf{v} = a\mathbf{i} + b\mathbf{j}$ es ortogonal a la recta $ax + by + c = 0$.
38. Muestre que el vector $\mathbf{u} = b\mathbf{i} - a\mathbf{j}$ es paralelo a la recta $ax + by + c = 0$.
39. Un triángulo tiene como vértices a $(1, 3)$, $(4, -2)$ y $(-3, 6)$. Encuentre el coseno de cada uno de sus ángulos.

40. Un triángulo tiene como vértices a (a_1, b_1), (a_2, b_2) y (a_3, b_3). Encuentre una fórmula para los cosenos de cada uno de sus ángulos.

★ 41. La *desigualdad de Cauchy-Schwarz* establece que para cualquier conjunto de número reales a_1, a_2, b_1 y b_2.

$$\left|\sum_{k=1}^{2} a_k b_k\right| \le \left(\sum_{k=1}^{2} a_k^2\right)^{1/2}\left(\sum_{k=1}^{2} b_k^2\right)^{1/2}$$

Use el producto escalar para demostrar esta fórmula. ¿En qué circunstancias esta desigualdad puede ser reemplazada por una igualdad?

★ 42. Demuestre que la distancia más corta entre un punto y una recta se mide a lo largo de una línea que pasa por el punto y es perpendicular a la recta.

43. Encuentre la distancia entre $P = (2, 3)$ y la recta que pasa por los puntos $Q = (-1, 7)$ y $R = (3, 5)$.

44. Encuentre la distancia entre $(3, 7)$ y la recta que contiene el vector $\mathbf{v} = 2\mathbf{i} - 3\mathbf{j}$ y pasa por el origen.

45. Sea A una matriz de 2×2 tal que cada columna es un vector unitario y las dos columnas son ortogonales. Demuestre que A es invertible y que $A^{-1} = A^t$. (A se denomina matriz *ortogonal*.)

3.3 Vectores en el espacio

Hemos visto que cualquier punto en un plano se puede representar como un par ordenado de números reales. Análogamente, cualquier punto en el espacio se puede representar por una *tríada ordenada* de números reales

$$(a, b, c) \tag{1}$$

$\mathbb{R}^3$ está compuesto de vectores de la forma (1). Para representar un punto en el espacio empezamos por escoger un punto en $\mathbb{R}^3$. Llamamos a este punto el *origen*, denotado 0. Luego dibujamos tres ejes mutuamente perpendiculares que llamamos *eje x*, *eje y* y *eje z*. Estos ejes se pueden seleccionar de varias maneras, pero la selección más común es con los ejes x y y dibujados horizontalmente, con el eje z vertical. En cada eje escogemos una dirección positiva y medimos la distancia a lo largo de ese eje como el número de unidades en esta dirección positiva medidas desde el origen.

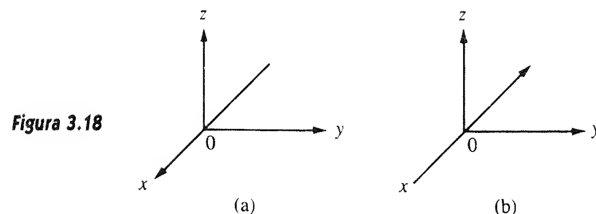


Figura 3.18

Los dos sistemas básicos para representar estos ejes se ven en la Figura 3.18. Si los ejes se ubican como en la Figura 3.18*a*, entonces se dice que el sistema es de *mano derecha*; si se colocan como en la Figura 3.18*b*, se dice que es de *mano izquierda*. En las figuras las flechas indican las direcciones positivas

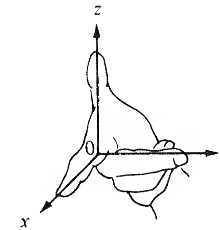


Figura 3.19

de los ejes. La justificación de estos términos es la que sigue: En un sistema de mano derecha, si ponemos la mano derecha de forma que el dedo índice apunte en la dirección positiva del eje x, mientras que el dedo medio apunte en la dirección positiva del eje y, entonces el pulgar apunta en la dirección positiva del eje z. Este concepto se ilustra en la Figura 3.19. Para un sistema de mano izquierda la misma regla se aplica para la mano izquierda. En lo que resta de este texto seguiremos la práctica común y dibujaremos los ejes coordenados usando un sistema de mano derecha.

Los tres ejes en nuestro sistema determinan tres *planos coordenados* que son llamados el plano xy, el xz y el yz. El plano xy contiene a los ejes x y y, y es simplemente el plano con el cual hemos estado tratando en la mayor parte de este libro. Los planos xz y yz se pueden considerar de manera análoga.

Habiendo construido nuestra estructura de ejes y planos coordenados, podemos describir cualquier punto P en $\mathbb{R}^3$ en forma única:

$$P = (x, y, z) \tag{2}$$

donde la primera coordenada x es la distancia del plano yz a P (medida en la dirección positiva del eje x y a lo largo de una línea paralela al eje x), la segunda coordenada y es la distancia del plano xz a P (medida en la dirección positiva del eje y y a lo largo de una línea paralela al eje y) y la tercera coordenada z es la distancia del plano xy a P (medida en la dirección positiva del eje z y a lo largo de una línea paralela al eje z).

En este sistema los tres planos coordenados dividen $\mathbb{R}^3$ en ocho *octantes* al igual que en $\mathbb{R}^2$ los dos ejes coordenados dividen el plano en cuatro cuadrantes. El primer octante es siempre aquél en el que las tres coordenadas son positivas.

El sistema coordenado así escogido se llama frecuentemente *sistema coordenado rectangular* o *sistema coordenado cartesiano*. Una vez que nos familiaricemos con la forma de describir un punto en este sistema, podremos generalizar varios conceptos del plano.

Teorema 1 Sean $P = (x_1, y_1, z_1)$ y $Q = (x_2, y_2, z_2)$ dos puntos en el espacio. Entonces la distancia $\overline{PQ}$ entre P y Q está dada por

$$\overline{PQ} = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2} \tag{3}$$

En el Problema 39 se pide demostrar este resultado.

Ejemplo 1 Calcule la distancia entre los puntos (3, −1, 6) y (−2, 3, 5).

Solución $\overline{PQ}=\sqrt{[3-(-2)]^2+(-1-3)^2+(6-5)^2}=\sqrt{42}$

En las Secciones 3.1 y 3.2 discutimos propiedades geométricas de los vectores en el plano. Debido a que los sistemas coordenados en $\mathbb{R}^2$ y $\mathbb{R}^3$ son muy similares, no es sorprendente que los vectores en $\mathbb{R}^2$ y $\mathbb{R}^3$, tengan estructuras muy similares. Discutiremos ahora el concepto de un vector en el espacio. El desarrollo de este tema seguirá estrechamente el desarrollo de las últimas dos secciones y por tanto se omitirán algunos detalles.

Sean P y Q dos puntos distintos en $\mathbb{R}^3$. Entonces el *segmento de recta dirigido* $\overrightarrow{PQ}$ es el segmento de recta que va de P a Q. Dos segmentos de recta dirigidos son *equivalentes* si tienen la misma magnitud y dirección. Un *vector* en $\mathbb{R}^3$ es el conjunto de todos los segmentos dirigidos equivalentes a un segmento dirigido dado y cualquier segmento dirigido $\overrightarrow{PQ}$ en ese conjunto se llama un *representante* del vector.

Hasta aquí las definiciones son idénticas. Por conveniencia elegimos P como el origen, de forma que el vector $\mathbf{v} = \overrightarrow{OQ}$ se pueda describir por las coordenadas (x, y, z) del punto Q. Entonces la magnitud de $\mathbf{v}=|\mathbf{v}|=\sqrt{x^2+y^2+z^2}$ (Teorema 1).

Ejemplo 2 Sea $\mathbf{v}=(1, 3, -2)$. Encuentre $|\mathbf{v}|$.

Solución $|\mathbf{v}|=\sqrt{1^2+3^2+(-2)}=\sqrt{14}$.

Sean $\mathbf{u}=(x_1, y_1, z_1)$ y $\mathbf{v}=(x_2, y_2, z_2)$ dos vectores y sea α un número real (escalar). Entonces definimos

$$\mathbf{u}+\mathbf{v}=(x_1+x_2,\ y_1+y_2,\ z_1+z_2)$$
y
$$\alpha\mathbf{u}=(\alpha x_1,\ \alpha y_1,\ \alpha z_1)$$

Esta es la misma definición de suma de vectores y multiplicación por un escalar que teníamos antes y se ilustra en la Figura 3.20.

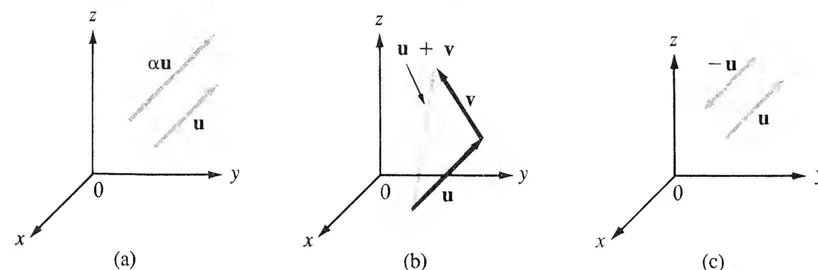


Figura 3.20

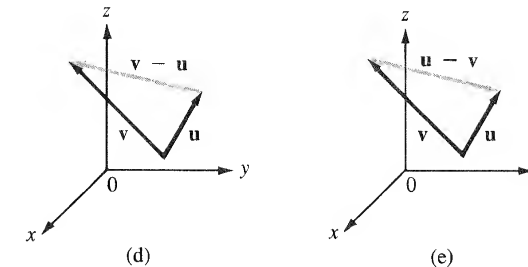


Figura 3.20

Un *vector unitario* $\mathbf{u}$ es un vector de magnitud 1. Si $\mathbf{v}$ es cualquier vector distinto de cero, entonces $\mathbf{u}=\mathbf{v}/|\mathbf{v}|$ es un vector unitario que tiene la misma dirección que $\mathbf{v}$.

Ejemplo 3 Encuentre un vector unitario que tenga la misma dirección que $\mathbf{v}=(2, 4, -3)$.

Solución Como $\mathbf{v}=\sqrt{2^2+4^2+(-3)^2}=\sqrt{29}$, tenemos que $\mathbf{u}=(2/\sqrt{29},\ 4/\sqrt{29},\ -3/\sqrt{29})$.

Ahora podemos definir formalmente la dirección de un vector en $\mathbb{R}^3$. No podemos definirla como el ángulo θ que forma el vector con la parte positiva del eje x pues, por ejemplo, si $0 < \theta < \pi/2$, entonces existe un *número infinito* de vectores que forman el ángulo θ con la parte positiva del eje x y todos ellos forman un cono (Figura 3.21).

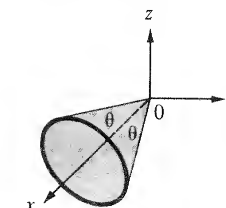


Figura 3.21

Definición 1 **Dirección en $\mathbb{R}^3$.** La *dirección* de un vector $\mathbf{v}$ distinto de cero en $\mathbb{R}^3$ se define como la dirección del vector unitario $\mathbf{u} = \mathbf{v}/|\mathbf{v}|$.

Observación. Podríamos haber definido la dirección de un vector $\mathbf{v}$ de $\mathbb{R}^2$ de esta forma. Si $\mathbf{u}=\mathbf{v}/|\mathbf{v}|$, entonces $\mathbf{u}=(\cos\theta, \operatorname{sen}\theta)$ donde θ es la dirección de $\mathbf{v}$.

Definiremos la dirección de un vector en términos de ciertos ángulos. Sea $\mathbf{v}$ el vector $\overrightarrow{OP}$ descrito en la Figura 3.22. Definimos α como el ángulo entre $\mathbf{v}$ y

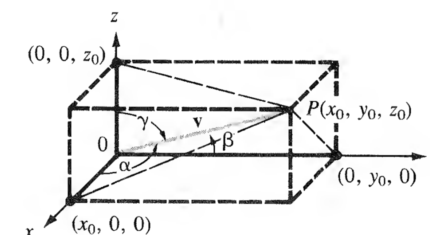


Figura 3.22

la parte positiva del eje x, β el ángulo entre $\mathbf{v}$ y la parte positiva del eje y y γ el ángulo entre $\mathbf{v}$ y la parte positiva del eje z. Los ángulos α, β y γ se conocen como los *ángulos directores* del vector $\mathbf{v}$. Entonces, de la Figura 3.22.

$$\cos\alpha = \frac{x_0}{|\mathbf{v}|} \qquad \cos\beta = \frac{y_0}{|\mathbf{v}|} \qquad \cos\gamma = \frac{z_0}{|\mathbf{v}|} \tag{4}$$

Si $\mathbf{v}$ es unitario entonces $|\mathbf{v}| = 1$ y

$$\cos\alpha = x_0 \qquad \cos\beta = y_0 \qquad \cos\gamma = z_0 \tag{5}$$

Por definición, cada uno de estos tres ángulos está en el intervalo $[0, \pi]$. Los cosenos de estos tres ángulos se denominan *cosenos directores* del vector $\mathbf{v}$. Notemos, de las Ecuaciones (4), que

$$\cos^2\alpha + \cos^2\beta + \cos^2\gamma = \frac{x_0^2 + y_0^2 + z_0^2}{|\mathbf{v}|^2} = \frac{x_0^2 + y_0^2 + z_0^2}{x_0^2 + y_0^2 + z_0^2} = 1 \tag{6}$$

Si α, β y γ son tres números cualesquiera entre 0 y π que satisfacen la condición (6), entonces determinan un vector único dado por $\mathbf{u} = (\cos\alpha, \cos\beta, \cos\gamma)$.

Observación. Si $\mathbf{v} = (a, b, c)$ y $\mathbf{v} \neq 1$, entonces los números a, b y c se conocen como los *números directores* del vector $\mathbf{v}$.

Ejemplo 4 Encuentre los cosenos directores del vector $\mathbf{v} = (4, -1, 6)$.

Solución La dirección de $\mathbf{v}$ es $\mathbf{v}/|\mathbf{v}| = \mathbf{v}/\sqrt{53} = (4/\sqrt{53}, -1/\sqrt{53}, 6/\sqrt{53})$. Entonces $\cos\alpha = 4/\sqrt{53} \approx 0.5494$, $\cos\beta = -1/\sqrt{53} \approx -0.1374$ y $\cos\gamma = 6/\sqrt{53} \approx 0.8242$. De aquí, usando una tabla de cosenos o una calculadora de bolsillo obtenemos $\alpha \approx 56.7° \approx 0.989$ rad, $\beta \approx 97.9° \approx 1.71$ rad y $\gamma = 34.5° \approx 0.602$ rad. El vector, junto con sus ángulos α, β y γ, aparece en la Figura 3.23.

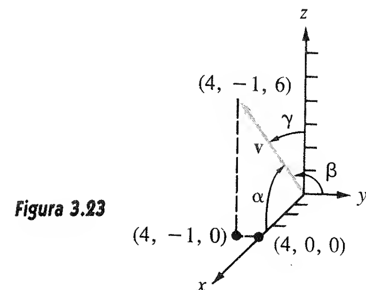


Figura 3.23

Ejemplo 5 Encuentre un vector $\mathbf{v}$ de magnitud 7 y cuyos cosenos directores sean $1/\sqrt{6}$, $1/\sqrt{3}$ y $1/\sqrt{2}$.

Sea $\mathbf{u} = (1/\sqrt{6}, 1/\sqrt{3}, 1/\sqrt{2})$. Entonces $\mathbf{u}$ es un vector unitario pues $|\mathbf{u}| = 1$. Así, la dirección de $\mathbf{v}$ está dada por $\mathbf{u}$ y $\mathbf{v} = |\mathbf{v}|\mathbf{u} = 7\mathbf{u} = (7/\sqrt{6}, 7/\sqrt{3}, 7/\sqrt{2})$.

Nota. Podemos resolver este problema porque $(1/\sqrt{6})^2 + (1/\sqrt{3})^2 + (1/\sqrt{2})^2 = 1$.

Es interesante notar que si $\mathbf{v}$, un vector en $\mathbb{R}^2$, se escribe $\mathbf{v} = (\cos\theta)\mathbf{i} + (\text{sen}\,\theta)\mathbf{j}$, donde θ es la dirección de $\mathbf{v}$, entonces $\cos\theta$ y $\text{sen}\,\theta$ son los cosenos directores de $\mathbf{v}$. Aquí $\alpha = \theta$ y definimos β como el ángulo que $\mathbf{v}$ forma con el eje y (vea Figura 3.24). Entonces $\beta = (\pi/2) - \alpha$ y así $\cos\beta = \cos(\pi/2 - \alpha) = \text{sen}\,\alpha$ y $\mathbf{v}$ se puede escribir en la forma de "cosenos directores":

$$\mathbf{v} = \cos\alpha\,\mathbf{i} + \cos\beta\,\mathbf{j}$$

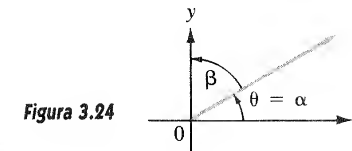


Figura 3.24

En la Sección 3.1 vimos que cualquier vector en el plano se puede escribir en términos de los vectores básicos $\mathbf{i}$ y $\mathbf{j}$. Para extender esta idea a $\mathbb{R}^3$ definimos

$$\mathbf{i} = (1, 0, 0) \qquad \mathbf{j} = (0, 1, 0) \qquad \mathbf{k} = (0, 0, 1) \tag{7}$$

Aquí, $\mathbf{i}$, $\mathbf{j}$ y $\mathbf{k}$ son vectores unitarios. El vector $\mathbf{i}$ está sobre el eje x, $\mathbf{j}$ en el eje y y $\mathbf{k}$ en el eje z (se representan en la Figura 3.25). Si $\mathbf{v} = (x, y, z)$ es un vector cualquiera en $\mathbb{R}^3$, entonces

$$\mathbf{v} = (x, y, z) = (x, 0, 0) + (0, y, 0) + (0, 0, z) = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$$

Esto es: *Cualquier vector* $\mathbf{v}$ *en* $\mathbb{R}^3$ *se puede escribir de forma única en términos de los vectores* $\mathbf{i}$, $\mathbf{j}$ *y* $\mathbf{k}$.

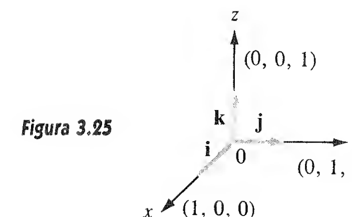


Figura 3.25

La definición de producto escalar en $\mathbb{R}^3$ es, desde luego, la definición que dimos en la Sección 1.5. Notemos que $\mathbf{i}\cdot\mathbf{i} = 1$, $\mathbf{j}\cdot\mathbf{j} = 1$, $\mathbf{k}\cdot\mathbf{k} = 1$, $\mathbf{i}\cdot\mathbf{j} = 0$, $\mathbf{j}\cdot\mathbf{k} = 0$, $\mathbf{i}\cdot\mathbf{k} = 0$.

Teorema 2 Si φ denota el menor ángulo positivo entre dos vectores **u** y **v** distintos de cero, tenemos que

$$\cos\varphi = \frac{\mathbf{u}\cdot\mathbf{v}}{|\mathbf{u}|\,|\mathbf{v}|} \tag{8}$$

Demostración La demostración es casi idéntica a la demostración del Teorema 3.2.1 y se deja como ejercicio (Problema 40).

Ejemplo 6 Calcule el coseno del ángulo entre $\mathbf{u}=3\mathbf{i}-\mathbf{j}+2\mathbf{k}$ y $\mathbf{v}=4\mathbf{i}+3\mathbf{j}-\mathbf{k}$.

Solución $\mathbf{u}\cdot\mathbf{v}=7$, $|\mathbf{u}|=\sqrt{14}$, $|\mathbf{v}|=\sqrt{26}$, de modo que $\cos\varphi = 7/\sqrt{(14)(26)} = 7/\sqrt{364} \approx 0.3669$ y $\varphi\approx 68.5°\approx 1.2$ rad.

Definición 2 ***Vectores paralelos y ortogonales.*** Dos vectores **u** y **v** distintos de cero son

i. *Paralelos* si el ángulo entre ellos es cero o bien π.
ii. *Ortogonales* (o *perpendiculares*) si el ángulo entre ellos es $\pi/2$.

Teorema 3

i. Si $\mathbf{u}\neq\mathbf{0}$, entonces **u** y **v** son paralelos si y sólo si $\mathbf{v}=\alpha\mathbf{u}$ para alguna constante $\alpha\neq 0$.
ii. Si **u** y **v** son distintos de cero, entonces **u** y **v** son ortogonales si y sólo si $\mathbf{u}\cdot\mathbf{v}=0$.

Demostración Nuevamente la demostración es fácil y se deja como ejercicio (Problema 41).

Ejemplo 7 Muestre que los vectores $\mathbf{u}=\mathbf{i}+3\mathbf{j}-4\mathbf{k}$ y $\mathbf{v}=-2\mathbf{i}-6\mathbf{j}+8\mathbf{k}$ son paralelos.

Solución $\mathbf{u}\cdot\mathbf{v} = -52$, $|\mathbf{u}| = \sqrt{26}$ y $|\mathbf{v}| = \sqrt{104} = 2\sqrt{26}$. Así, $\mathbf{u}\cdot\mathbf{v}/|\mathbf{u}|\,|\mathbf{v}| = -52/(\sqrt{26}\cdot 2\sqrt{26}) = -1$, de manera que $\cos\theta = -1$, $\theta = \pi$ y **u** y **v** son paralelos (pero de direcciones opuestas). Otra forma de ver esto es notando que $\mathbf{v} = -2\mathbf{u}$ y, por el Teorema 3, **u** y **v** son paralelos (Figura 3.26).

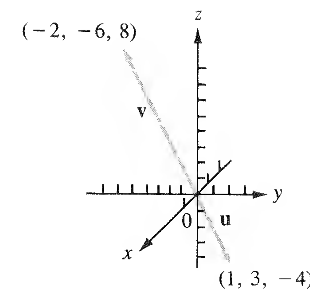


Figura 3.26

Ejemplo 8 Encuentre un número α tal que $\mathbf{u} = 8\mathbf{i} - 2\mathbf{j} + 4\mathbf{k}$ y $\mathbf{v} = 2\mathbf{i} + 3\mathbf{j} + \alpha\mathbf{k}$ sean ortogonales.

Solución Debemos mostrar que $0=\mathbf{u}\cdot\mathbf{v}=10+4\alpha$ de donde $\alpha=-\frac{5}{2}$. Los vectores **u** y **v** están dibujados en la Figura 3.27.

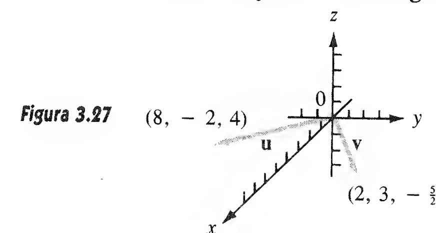


Figura 3.27

Volvamos ahora a la definición de la proyección de un vector sobre otro. Primero estableceremos un teorema análogo al Teorema 3.2.4 (y que tiene idéntica demostración).

Teorema 4 Sea **v** un vector distinto de cero. Entonces, para cualquier otro vector **u**,

$$\mathbf{w} = \mathbf{u} - \frac{\mathbf{u}\cdot\mathbf{v}}{|\mathbf{v}|^2}\mathbf{v}$$

es ortogonal a **v**.

Definición 3 ***Proyección.*** Sean **u** y **v** dos vectores distintos de cero. Entonces la *proyección de* **u** *sobre* **v**, denotada $\text{proy}_{\mathbf{v}}\mathbf{u}$, se define como

$$\text{proy}_{\mathbf{v}}\,\mathbf{u} = \frac{\mathbf{u}\cdot\mathbf{v}}{|\mathbf{v}|^2}\mathbf{v} \tag{9}$$

La *componente* de **u** en la dirección **v** está dada por $(\mathbf{u}\cdot\mathbf{v})/|\mathbf{v}|$.

Ejemplo 9 Sean $\mathbf{u}=2\mathbf{i}+3\mathbf{j}+\mathbf{k}$ y $\mathbf{v}=\mathbf{i}+2\mathbf{j}-6\mathbf{k}$. Encuentre $\text{proy}_{\mathbf{v}}\mathbf{u}$.

Solución Aquí $(\mathbf{u}\cdot\mathbf{v})/|\mathbf{v}|^2=2/41$ y $\text{proy}_{\mathbf{v}}\mathbf{u}=\frac{2}{41}\mathbf{i}+\frac{4}{41}\mathbf{j}-\frac{12}{41}\mathbf{k}$. La componente de $\mathbf{u}$ en la dirección $\mathbf{v}$ es $(\mathbf{u}\cdot\mathbf{v})/|\mathbf{v}|=2/\sqrt{41}$.

Notemos que, como en el caso del plano, $\text{proy}_{\mathbf{v}}\mathbf{u}$ es un vector que tiene la misma dirección que $\mathbf{v}$ si $\mathbf{u}\cdot\mathbf{v}>0$ y dirección opuesta a la de $\mathbf{v}$ si $\mathbf{u}\cdot\mathbf{v}<0$.

Problemas 3.3

En los Problemas del 1 al 3 encuentre la distancia entre los dos puntos.

1. (3, −4, 3); (3, 2, 5)
2. (3, −4, 7); (3, −4, 9)
3. (−2, 1, 3); (4, 1, 3)

En los Problemas del 4 al 17 determine la magnitud y los cosenos directores del vector dado.

4. $\mathbf{v}=3\mathbf{j}$
5. $\mathbf{v}=-3\mathbf{i}$
6. $\mathbf{v}=4\mathbf{i}-\mathbf{j}$
7. $\mathbf{v}=\mathbf{i}+2\mathbf{k}$
8. $\mathbf{v}=\mathbf{i}-\mathbf{j}+\mathbf{k}$
9. $\mathbf{v}=\mathbf{i}+\mathbf{j}-\mathbf{k}$
10. $\mathbf{v}=-\mathbf{i}+\mathbf{j}+\mathbf{k}$
11. $\mathbf{v}=\mathbf{i}-\mathbf{j}-\mathbf{k}$
12. $\mathbf{v}=-\mathbf{i}+\mathbf{j}-\mathbf{k}$
13. $\mathbf{v}=-\mathbf{i}-\mathbf{j}+\mathbf{k}$
14. $\mathbf{v}=-\mathbf{i}-\mathbf{j}-\mathbf{k}$
15. $\mathbf{v}=2\mathbf{i}+5\mathbf{j}-7\mathbf{k}$
16. $\mathbf{v}=-3\mathbf{i}-3\mathbf{j}+8\mathbf{k}$
17. $\mathbf{v}=-2\mathbf{i}-3\mathbf{j}-4\mathbf{k}$
18. Los tres ángulos directores de cierto vector unitario son los mismos y están entre 0 y $\pi/2$. ¿Cuál es el vector?
19. Encuentre un vector de magnitud 12 que tenga la misma dirección que el vector del Problema 18.
20. Muestre que no existe ningún vector unitario cuyos ángulos directores sean $\pi/6$, $\pi/3$ y $\pi/4$.
21. Sean $P=(2, 1, 4)$ y $Q=(3, -2, 8)$. Encuentre un vector unitario en la dirección $\overrightarrow{PQ}$.
22. Sean $P=(-3, 1, 7)$ y $Q=(8, 1, 7)$. Encuentre un vector unitario cuya dirección sea la opuesta a la de $\overrightarrow{PQ}$.
23. En el Problema 22 encuentre todos los puntos R tales que $\overrightarrow{PR}\perp\overrightarrow{PQ}$.
24. ★ Muestre que el conjunto de puntos que satisfacen la condición del Problema 23 y la condición $|\overrightarrow{PR}|=1$ forman un círculo.
25. Si $\mathbf{u}$ y $\mathbf{v}$ están en $\mathbb{R}^3$, muestre que $|\mathbf{u}+\mathbf{v}|\le|\mathbf{u}|+|\mathbf{v}|$.
26. ¿En qué circunstancias la desigualdad del Problema 25 se puede sustituir por una igualdad?

En los Problemas 27 al 38 sean $\mathbf{u}=2\mathbf{i}-3\mathbf{j}+4\mathbf{k}$, $\mathbf{v}=-2\mathbf{i}-3\mathbf{j}+5\mathbf{k}$, $\mathbf{w}=\mathbf{i}-7\mathbf{j}+3\mathbf{k}$ y $\mathbf{t}=3\mathbf{i}+4\mathbf{j}+5\mathbf{k}$.

27. Calcule $\mathbf{u}+\mathbf{v}$.
28. Calcule $2\mathbf{u}-3\mathbf{v}$.
29. Calcule $\mathbf{t}+3\mathbf{w}-\mathbf{v}$.
30. Calcule $2\mathbf{u}-7\mathbf{w}+5\mathbf{v}$.
31. Calcule $2\mathbf{v}+7\mathbf{t}-\mathbf{w}$.
32. Calcule $\mathbf{u}\cdot\mathbf{v}$.
33. Calcule $|\mathbf{w}|$.
34. Calcule $\mathbf{u}\cdot\mathbf{w}-\mathbf{w}\cdot\mathbf{t}$.
35. Calcule el ángulo entre $\mathbf{u}$ y $\mathbf{w}$.
36. Calcule el ángulo entre $\mathbf{t}$ y $\mathbf{w}$.
37. Calcule $\text{proy}_{\mathbf{u}}\mathbf{v}$.
38. Calcule $\text{proy}_{\mathbf{t}}\mathbf{w}$.
39. Demuestre el Teorema 1. [*Sugerencia:* Usar el teorema de Pitágoras dos veces en la Figura 3.28.]

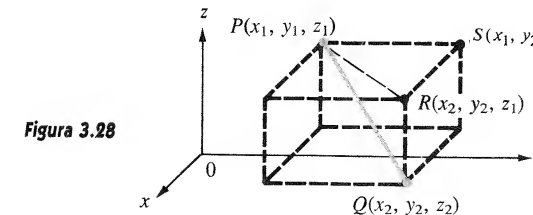


Figura 3.28

40. Demuestre el Teorema 2.
41. Demuestre el Teorema 3.
42. Demuestre el Teorema 4.

3.4 El producto vectorial (o cruz) de dos vectores

Hasta aquí el único producto de vectores que hemos considerado es el producto escalar o producto punto. Definimos ahora un nuevo producto llamado *producto vectorial* (o *producto cruz*)* que sólo está definido en $\mathbb{R}^3$.

Definición 1 **Producto vectorial.** Sea $\mathbf{u}=a_1\mathbf{i}+b_1\mathbf{j}+c_1\mathbf{k}$ y $\mathbf{v}=a_2\mathbf{i}+b_2\mathbf{j}+c_2\mathbf{k}$. Entonces el *producto cruz* (o *vectorial*) de $\mathbf{u}$ y $\mathbf{v}$, denotado $\mathbf{u}\times\mathbf{v}$, es un nuevo vector definido por

$$\mathbf{u}\times\mathbf{v}=(b_1c_2-c_1b_2)\mathbf{i}+(c_1a_2-a_1c_2)\mathbf{j}+(a_1b_2-b_1a_2)\mathbf{k} \quad (1)$$

Observe que el resultado del producto cruz es un vector, mientras el resultado del producto escalar es un escalar.

Parecería que el producto cruz ha sido definido de una manera un tanto arbitraria. Existen obviamente varias maneras de definir un producto vectorial. ¿Por qué se escogió esta definición? Responderemos a esta pregunta en esta sección demostrando algunas de las propiedades del producto cruz e ilustrando algunos de sus usos.

Ejemplo 1 Sea $\mathbf{u}=\mathbf{i}-\mathbf{j}+2\mathbf{k}$ y $\mathbf{v}=2\mathbf{i}+3\mathbf{j}-4\mathbf{k}$. Calcule $\mathbf{w}=\mathbf{u}\times\mathbf{v}$.

* *Nota histórica:* El producto cruz fue definido por Hamilton en uno de los artículos en que discutía los cuaterniones, publicados en *Philosophical Magazine* entre los años 1844 y 1850.

Solución Usando la fórmula (1),

$$\mathbf{w}=[(-1)(-4)-(2)(3)]\mathbf{i}+[(2)(2)-(1)(-4)]\mathbf{j}+[(1)(3)-(-1)(2)]\mathbf{k}$$
$$=-2\mathbf{i}+8\mathbf{j}+5\mathbf{k}$$

Nota. En este ejemplo $\mathbf{u}\cdot\mathbf{w}=(\mathbf{i}-\mathbf{j}+2\mathbf{k})\cdot(-2\mathbf{i}+8\mathbf{j}+5\mathbf{k})=-2-8+10=0$. Análogamente $\mathbf{v}\cdot\mathbf{w}=0$. Esto es, $\mathbf{u}\times\mathbf{v}$ es ortogonal tanto a $\mathbf{u}$ como a $\mathbf{v}$. Como veremos en breve, el producto cruz de $\mathbf{u}$ y $\mathbf{v}$ siempre es ortogonal a $\mathbf{u}$ y a $\mathbf{v}$.

Antes de continuar nuestra discusión de los usos del producto cruz observemos que hay una forma fácil de calcular $\mathbf{u}\times\mathbf{v}$ con el uso de determinantes.

Teorema 1

$$\mathbf{u}\times\mathbf{v}=\begin{vmatrix}\mathbf{i}&\mathbf{j}&\mathbf{k}\\a_1&b_1&c_1\\a_2&b_2&c_2\end{vmatrix}^{*}$$

Demostración

$$\begin{vmatrix}\mathbf{i}&\mathbf{j}&\mathbf{k}\\a_1&b_1&c_1\\a_2&b_2&c_2\end{vmatrix}=\mathbf{i}\begin{vmatrix}b_1&c_1\\b_2&c_2\end{vmatrix}-\mathbf{j}\begin{vmatrix}a_1&c_1\\a_2&c_2\end{vmatrix}+\mathbf{k}\begin{vmatrix}a_1&b_1\\a_2&b_2\end{vmatrix}$$
$$=(b_1c_2-c_1b_2)\mathbf{i}+(c_1a_2-a_1c_2)\mathbf{j}+(a_1b_2-b_1a_2)\mathbf{k}$$

lo que es igual a $\mathbf{u}\times\mathbf{v}$ de acuerdo a la Definición 1. ■

Ejemplo 2 Calcule $\mathbf{u}\times\mathbf{v}$, donde $\mathbf{u}=2\mathbf{i}+4\mathbf{j}-5\mathbf{k}$ y $\mathbf{v}=-3\mathbf{i}-2\mathbf{j}+\mathbf{k}$.

Solución

$$\mathbf{u}\times\mathbf{v}=\begin{vmatrix}\mathbf{i}&\mathbf{j}&\mathbf{k}\\2&4&-5\\-3&-2&1\end{vmatrix}=(4-10)\mathbf{i}-(2-15)\mathbf{j}+(-4+12)\mathbf{k}$$
$$=-6\mathbf{i}+13\mathbf{j}+8\mathbf{k}$$

El siguiente teorema resume algunas de las propiedades del producto cruz. Su demostración se deja como ejercicio (vea Problemas del 32 al 35).

Teorema 2 Sean $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ vectores en $\mathbb{R}^3$ y sea α un escalar. Entonces:

i. $\mathbf{u}\times\mathbf{0}=\mathbf{0}\times\mathbf{u}=\mathbf{0}$.
ii. $\mathbf{u}\times\mathbf{v}=-(\mathbf{v}\times\mathbf{u})$ **(propiedad anticonmutativa del producto vectorial).**

* Este no es realmente un determinante pues **i**, **j** y **k** no son números. Sin embargo, usando notación de determinantes, el Teorema 1 nos ayuda a recordar cómo se calcula un producto cruz.

iii. $(\alpha\mathbf{u})\times\mathbf{v}=\alpha(\mathbf{u}\times\mathbf{v})$.
iv. $\mathbf{u}\times(\mathbf{v}+\mathbf{w})=(\mathbf{u}\times\mathbf{v})+(\mathbf{u}\times\mathbf{w})$ **(propiedad distributiva del producto vectorial).**
v. $(\mathbf{u}\times\mathbf{v})\cdot\mathbf{w}=\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w})$. (Esto es llamado el *triple producto escalar* de $\mathbf{u}$, $\mathbf{v}$, y $\mathbf{w}$.)
vi. $\mathbf{u}\cdot(\mathbf{u}\times\mathbf{v})=\mathbf{v}\cdot(\mathbf{u}\times\mathbf{v})=0$. (Esto es, $\mathbf{u}\times\mathbf{v}$ es ortogonal tanto a $\mathbf{u}$ como a $\mathbf{v}$.)
vii. Si $\mathbf{u}$ y $\mathbf{v}$ son paralelos entonces $\mathbf{u}\times\mathbf{v}=\mathbf{0}$.

La parte (*vi*) es la más comúnmente usada de este teorema. La reformularemos a continuación:

El producto cruz $\mathbf{u}\times\mathbf{v}$ es ortogonal a $\mathbf{u}$ y $\mathbf{v}$.

Sabemos que $\mathbf{u}\times\mathbf{v}$ es un vector ortogonal tanto a $\mathbf{u}$ como a $\mathbf{v}$. El siguiente resultado nos da su magnitud.

Teorema 3 Si φ es el ángulo entre $\mathbf{u}$ y $\mathbf{v}$, entonces

$$|\mathbf{u}\times\mathbf{v}|=|\mathbf{u}|\,|\mathbf{v}|\operatorname{sen}\varphi \qquad (2)$$

Demostración Es fácil mostrar (comparando componentes) que $|\mathbf{u}\times\mathbf{v}|^2=|\mathbf{u}|^2|\mathbf{v}|^2-(\mathbf{u}\cdot\mathbf{v})^2$ (Problema 31). Entonces, como $(\mathbf{u}\cdot\mathbf{v})^2=|\mathbf{u}|^2|\mathbf{v}|^2\cos^2\varphi$ (del Teorema 3.3.2),

$$|\mathbf{u}\times\mathbf{v}|^2=|\mathbf{u}|^2|\mathbf{v}|^2-|\mathbf{u}|^2|\mathbf{v}|^2\cos^2\varphi=|\mathbf{u}|^2|\mathbf{v}|^2(1-\cos^2\theta)$$
$$=|\mathbf{u}|^2|\mathbf{v}|^2\operatorname{sen}^2\varphi$$

y el teorema queda demostrado tomando la raíz cuadrada de ambos lados. ■

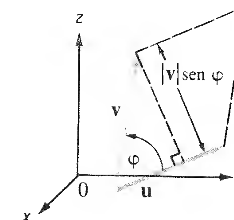


Figura 3.29

Existe una interpretación geométrica interesante del Teorema 3. Los vectores $\mathbf{u}$ y $\mathbf{v}$ están dibujados en la Figura 3.29 y se los puede imaginar como los dos lados adyacentes de un paralelogramo. Entonces, de la geometría elemental, vemos que

$$\text{Área del paralelogramo} = |\mathbf{u}|\,|\mathbf{v}| \operatorname{sen} \varphi = |\mathbf{u} \times \mathbf{v}| \qquad (3)$$

Ejemplo 3 Encuentre el área del paralelogramo con vértices consecutivos en $P=(1, 3, -2)$, $Q=(2, 1, 4)$ y $R=(-3, 1, 6)$.

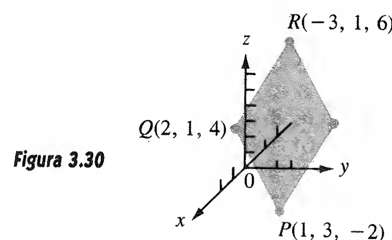


Figura 3.30

Solución El paralelogramo está dibujado en la Figura 3.30. Tenemos

$$\text{Área} = |\overrightarrow{PQ} \times \overrightarrow{QR}| = |(\mathbf{i} - 2\mathbf{j} + 6\mathbf{k}) \times (-5\mathbf{i} + 2\mathbf{k})|$$

$$= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & -2 & 6 \\ -5 & 0 & 2 \end{vmatrix} = |-4\mathbf{i} - 32\mathbf{j} - 10\mathbf{k}| = \sqrt{1140} \text{ unidades cuadradas}$$

Interpretación geométrica de determinantes 2 × 2

Podemos usar la discusión anterior para dar una interpretación geométrica del determinante. Sea A una matriz de 2×2 y sean $\mathbf{u}$ y $\mathbf{v}$ dos vectores con dos

Sean $\mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$ y $\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$. Estos vectores están dados en la Figura 3.31. El *área generada* por $\mathbf{u}$ *y* $\mathbf{v}$ se define como el área del paralelogramo dado en la figura. Podemos considerar a $\mathbf{u}$ y $\mathbf{v}$ como vectores de $\mathbb{R}^3$ en el plano xy.

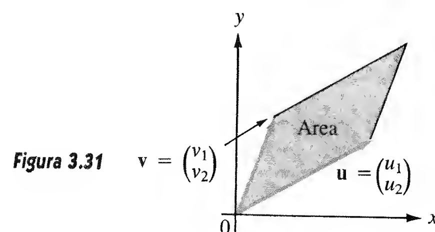


Figura 3.31

Entonces $\mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \\ 0 \end{pmatrix}$, $\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ 0 \end{pmatrix}$ y por tanto

$$\text{área generada por } \mathbf{u} \text{ y } \mathbf{v} = |\mathbf{u} \times \mathbf{v}| = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ u_1 & u_2 & 0 \\ v_1 & v_2 & 0 \end{vmatrix}$$

$$= |(u_1 v_2 - u_2 v_1)\mathbf{k}| = |u_1 v_2 - u_2 v_1|^*$$

Ahora sean $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, $\mathbf{u}' = A\mathbf{u}$ y $\mathbf{v}' = A\mathbf{v}$. Entonces

$$\mathbf{u}' = \begin{pmatrix} a_{11}u_1 + a_{12}u_2 \\ a_{21}u_1 + a_{22}u_2 \end{pmatrix} \text{ y } \mathbf{v}' = \begin{pmatrix} a_{11}v_1 + a_{12}v_2 \\ a_{21}v_1 + a_{22}v_2 \end{pmatrix}.$$

¿Cuál es el área generada por $\mathbf{u}'$ y $\mathbf{v}'$? Siguiendo los pasos anteriores,

$$\text{Área generada por } \mathbf{u}' \text{ y } \mathbf{v}' = |\mathbf{u}' \times \mathbf{v}'| = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_{11}u_1 + a_{12}u_2 & a_{21}u_1 + a_{22}u_2 & 0 \\ a_{11}v_1 + a_{12}v_2 & a_{21}v_1 + a_{22}v_2 & 0 \end{vmatrix}$$

$$= |(a_{11}u_1 + a_{12}u_2)(a_{21}v_1 + a_{22}v_2) - (a_{21}u_1 + a_{22}u_2)(a_{11}v_1 + a_{12}v_2)|$$

Por álgebra elemental podemos verificar que la última expresión es igual a

$$|(a_{11}a_{22} - a_{12}a_{21})(u_1 v_2 - u_2 v_1)| = \pm\det A \text{ (área generada por } \mathbf{u} \text{ y } \mathbf{v})$$

Así (en este contexto): *El determinante tiene el efecto de multiplicar el área.* En el Problema 41 se pide mostrar que, en cierto sentido, un determinante de 3×3 tiene el efecto de multiplicar el volumen.

Problemas 3.4

En los Problemas del 1 *al* 20 *determine el producto cruz* $\mathbf{u} \times \mathbf{v}$.

1. $\mathbf{u} = \mathbf{i} - 2\mathbf{j}$; $\mathbf{v} = 3\mathbf{k}$
2. $\mathbf{u} = 3\mathbf{i} - 7\mathbf{j}$; $\mathbf{v} = \mathbf{i} + \mathbf{k}$
3. $\mathbf{u} = \mathbf{i} - \mathbf{j}$; $\mathbf{v} = \mathbf{j} + \mathbf{k}$
4. $\mathbf{u} = -7\mathbf{k}$; $\mathbf{v} = \mathbf{j} + 2\mathbf{k}$
5. $\mathbf{u} = -2\mathbf{i} + 3\mathbf{j}$; $\mathbf{v} = 7\mathbf{i} + 4\mathbf{k}$
6. $\mathbf{u} = a\mathbf{i} + b\mathbf{j}$; $\mathbf{v} = c\mathbf{i} + d\mathbf{j}$
7. $\mathbf{u} = a\mathbf{i} + b\mathbf{k}$; $\mathbf{v} = c\mathbf{i} + d\mathbf{k}$
8. $\mathbf{u} = a\mathbf{j} + b\mathbf{k}$; $\mathbf{v} = c\mathbf{i} + d\mathbf{k}$
9. $\mathbf{u} = 2\mathbf{i} - 3\mathbf{j} + \mathbf{k}$; $\mathbf{v} = \mathbf{i} + 2\mathbf{j} + \mathbf{k}$
10. $\mathbf{u} = 3\mathbf{i} - 4\mathbf{j} + 2\mathbf{k}$; $\mathbf{v} = 6\mathbf{i} - 3\mathbf{j} + 5\mathbf{k}$
11. $\mathbf{u} = -3\mathbf{i} - 2\mathbf{j} + \mathbf{k}$; $\mathbf{v} = 6\mathbf{i} + 4\mathbf{j} - 2\mathbf{k}$
12. $\mathbf{u} = \mathbf{i} + 7\mathbf{j} - 3\mathbf{k}$; $\mathbf{v} = -\mathbf{i} - 7\mathbf{j} + 3\mathbf{k}$
13. $\mathbf{u} = \mathbf{i} - 7\mathbf{j} - 3\mathbf{k}$; $\mathbf{v} = -\mathbf{i} + 7\mathbf{j} - 3\mathbf{k}$
14. $\mathbf{u} = 2\mathbf{i} - 3\mathbf{j} + 5\mathbf{k}$; $\mathbf{v} = 3\mathbf{i} - \mathbf{j} - \mathbf{k}$
15. $\mathbf{u} = 10\mathbf{i} + 7\mathbf{j} - 3\mathbf{k}$; $\mathbf{v} = -3\mathbf{i} + 4\mathbf{j} - 3\mathbf{k}$
16. $\mathbf{u} = 2\mathbf{i} + 4\mathbf{j} - 6\mathbf{k}$; $\mathbf{v} = -\mathbf{i} - \mathbf{j} + 3\mathbf{k}$
17. $\mathbf{u} = 2\mathbf{i} - \mathbf{j} + \mathbf{k}$; $\mathbf{v} = 4\mathbf{i} + 2\mathbf{j} + 2\mathbf{k}$
18. $\mathbf{u} = 3\mathbf{i} - \mathbf{j} + 8\mathbf{k}$; $\mathbf{v} = \mathbf{i} + \mathbf{j} - 4\mathbf{k}$
19. $\mathbf{u} = a\mathbf{i} + a\mathbf{j} + a\mathbf{k}$; $\mathbf{v} = b\mathbf{i} + b\mathbf{j} + b\mathbf{k}$
20. $\mathbf{u} = a\mathbf{i} + b\mathbf{j} + c\mathbf{k}$; $\mathbf{v} = a\mathbf{i} + b\mathbf{j} - c\mathbf{k}$
21. Halle dos vectores unitarios ortogonales a $\mathbf{u} = 2\mathbf{i} - 3\mathbf{j}$ y a $\mathbf{v} = 4\mathbf{j} + 3\mathbf{k}$.
22. Encuentre dos vectores unitarios ortogonales a $\mathbf{u} = \mathbf{i} + \mathbf{j} + \mathbf{k}$ y a $\mathbf{v} = \mathbf{i} - \mathbf{j} - \mathbf{k}$.
23. Utilice el producto cruz para encontrar el seno del ángulo φ entre los vectores $\mathbf{u} = 2\mathbf{i} + \mathbf{j} - \mathbf{k}$ y $\mathbf{v} = -3\mathbf{i} - 2\mathbf{j} + 4\mathbf{k}$.
24. Utilice el producto escalar para calcular el coseno del ángulo φ entre los vectores del Problema 23. Luego muestre que para los valores calculados se cumple $\operatorname{sen}^2\varphi + \cos^2\varphi = 1$.

---

* Notemos que éste es el valor absoluto de $\det \begin{pmatrix} u_1 & v_1 \\ u_2 & v_2 \end{pmatrix}$.

En los Problemas del 25 al 30 encuentre el área del paralelogramo con los vértices adyacentes dados.

25. $(1, -2, 3)$; $(2, 0, 1)$; $(0, 4, 0)$
26. $(-2, 1, 1)$; $(2, 2, 3)$; $(-1, -2, 4)$
27. $(-2, 1, 0)$; $(1, 4, 2)$; $(-3, 1, 5)$
28. $(7, -2, -3)$; $(-4, 1, 6)$; $(5, -2, 3)$
29. $(a, 0, 0)$; $(0, b, 0)$; $(0, 0, c)$
30. $(a, b, 0)$; $(a, 0, b)$; $(0, a, b)$
31. Muestre que $|\mathbf{u} \times \mathbf{v}|^2 = |\mathbf{u}|^2|\mathbf{v}|^2 - (\mathbf{u} \cdot \mathbf{v})^2$. [*Sugerencia:* Desarrolle en términos de sus componentes.]
32. Use las Propiedades 1, 4, 2 y 3 (en ese orden) en la Sección 2.2 para comprobar las partes (*i*), (*ii*); (*iii*) y (*iv*) del Teorema 2.
33. Pruebe el Teorema 2(*v*) escribiendo en forma completa los componentes de cada lado de la igualdad.
34. Demuestre el Teorema 2(*vi*). [*Sugerencia:* Utilice las partes (*ii*) y (*v*) y el hecho de que el producto escalar es conmutativo para mostrar que $\mathbf{u} \cdot (\mathbf{u} \times \mathbf{v}) = -\mathbf{u} \cdot (\mathbf{u} \times \mathbf{v})$.]
35. Demuestre el Teorema 2(*vii*). [*Sugerencia:* Use el Teorema 3.33 y la Propiedad 6.]

$$\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$$

★ 37. Sean **u**, **v** y **w** tres vectores que no están en el mismo plano. Tales vectores forman los lados de un paralelepípedo en el espacio (Figura 3.32). Demuestre que el volumen del paralelepípedo está dado por $V = |(\mathbf{u} \times \mathbf{v}) \cdot \mathbf{w}|$.* [*Sugerencia:* El área de la base es $|\mathbf{u} \times \mathbf{v}|$.]

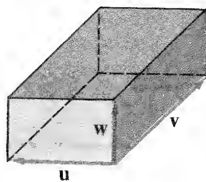


Figura 3.32

38. Calcule el volumen del paralelepípedo determinado por los vectores $\mathbf{u} = 2\mathbf{i} - \mathbf{j} + \mathbf{k}$, $\mathbf{v} = 3\mathbf{i} + 2\mathbf{j} - 2\mathbf{k}$ y $\mathbf{w} = 3\mathbf{i} + 2\mathbf{j}$.
39. Calcule el volumen del paralelepípedo determinado por los vectores $\mathbf{i} - \mathbf{j}$, $3\mathbf{i} + 2\mathbf{k}$, $-7\mathbf{j} + 3\mathbf{k}$.
40. Calcule el volumen del paralelepípedo determinado por los vectores $\overrightarrow{PQ}$, $\overrightarrow{PR}$ y $\overrightarrow{PS}$, donde $P = (2, 1, -1)$, $Q = (-3, 1, 4)$, $R = (-1, 0, 2)$ y $S = (-3, -1, 5)$.

★★ 41. El volumen generado por tres vectores **u**, **v** y **w** de $\mathbb{R}^3$ se define como el volumen del paralelepípedo cuyos lados son **u**, **v** y **w** (Figura 3.32). Sea A una matriz de 3×3 y sean $\mathbf{u}_1 = A\mathbf{u}$, $\mathbf{v}_1 = A\mathbf{v}$ y $\mathbf{w}_1 = A\mathbf{w}$. Muestre que:

Volumen generado por $\mathbf{u}_1$, $\mathbf{v}_1$ y $\mathbf{w}_1$ $= (\pm \det A)$(volumen generado por **u**, **v** y **w**)

* Esto significa que el volumen del paralelogramo está dado por

$$V = \left| \det \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} \right|$$

Esto muestra que al igual que el determinante de una matriz de 2×2 multiplica el área, el determinante de una matriz de 3×3 multiplica el volumen.

42. Sean $A = \begin{pmatrix} 2 & 3 & 1 \\ 4 & -1 & 5 \\ 1 & 0 & 6 \end{pmatrix}$, $\mathbf{u} = \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}$, $\mathbf{v} = \begin{pmatrix} 1 \\ 0 \\ 4 \end{pmatrix}$ y $\mathbf{w} = \begin{pmatrix} -1 \\ 3 \\ 2 \end{pmatrix}$.
    a. Calcule el volumen generado por **u**, **v** y **w**.
    b. Calcule el volumen generado por $A\mathbf{u}$, $A\mathbf{v}$ y $A\mathbf{w}$.
    c. Calcule el determinante de A.
    d. Muestre que [volumen en la parte (*b*)] = ($\pm$det A) $\times$ [volumen en la parte (*a*)].
43. El **triple producto cruz (o vectorial)** de tres vectores en $\mathbb{R}^3$ se define como $\mathbf{u} \times (\mathbf{v} \times \mathbf{w})$. Demuestre que

$$\mathbf{u} \times (\mathbf{v} \times \mathbf{w}) = (\mathbf{u} \cdot \mathbf{w})\mathbf{v} - (\mathbf{u} \cdot \mathbf{v})\mathbf{w}.$$

3.5 Rectas y planos en el espacio

En el plano $\mathbb{R}^2$ podemos encontrar la ecuación de una recta dados dos puntos de la recta o un punto y la pendiente de la recta. En $\mathbb{R}^3$, nuestra intuición nos dice que las ideas básicas son las mismas. Como dos puntos determinan una recta deberíamos ser capaces de calcular la ecuación de una recta en el espacio si conocemos dos puntos en ella. De otra forma si conocemos un punto y la dirección de la recta debemos también ser capaces de encontrar su ecuación.

Empezamos con dos puntos $P = (x_1, y_1, z_1)$ y $Q = (x_2, y_2, z_2)$ en una recta L. Un vector paralelo a L es un vector con representante $\overrightarrow{PQ}$. Por tanto,

$$\mathbf{v} = (x_2 - x_1)\mathbf{i} + (y_2 - y_1)\mathbf{j} + (z_2 - z_1)\mathbf{k} \tag{1}$$

es un vector paralelo a L. Ahora, sea $R = (x, y, z)$ otro punto en la recta. Entonces $\overrightarrow{PR}$ es paralelo a $\overrightarrow{PQ}$, que a su vez es paralelo a **v** y así, por el Teorema 3.3.3

$$\overrightarrow{PR} = t\mathbf{v} \tag{2}$$

para algún número real t. De la Figura 3.33 tenemos (en cada uno de los tres casos posibles)

$$\overrightarrow{OR} = \overrightarrow{OP} + \overrightarrow{PR} \tag{3}$$

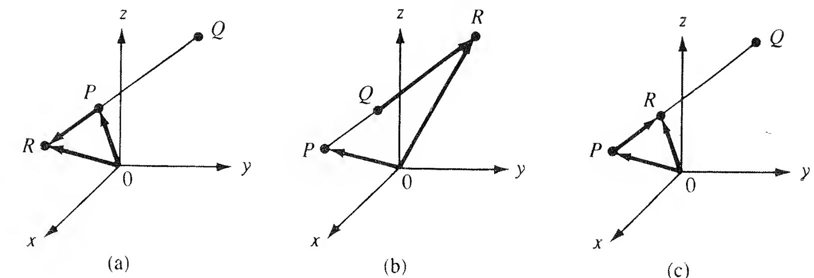


Figura 3.33

Y, combinando (2) y (3) obtenemos

$$\overrightarrow{PR} = \overrightarrow{OR} - \overrightarrow{OP} = t\mathbf{v}$$

o bien

$$\overrightarrow{OR} = \overrightarrow{OP} + t\mathbf{v} \qquad (4)$$

Ecuación vectorial de una recta

La Ecuación (4) se denomina *ecuación vectorial* de la recta L. Si R está en L, entonces (4) es satisfecha para algún número real t. Inversamente, si (4) es satisfecha, entonces, regresando sobre nuestros pasos, vemos que $\overrightarrow{PR}$ es paralela a $\mathbf{v}$, lo que significa que R está en L.

Si desarrollamos en componentes la Ecuación (4) obtenemos

$$x\mathbf{i} + y\mathbf{j} + z\mathbf{k} = x_1\mathbf{i} + y_1\mathbf{j} + z_1\mathbf{k} + t(x_2 - x_1)\mathbf{i} + t(y_2 - y_1)\mathbf{j} + t(z_2 - z_1)\mathbf{k}$$

o bien

$$x = x_1 + t(x_2 - x_1)$$
$$y = y_1 + t(y_2 - y_1) \qquad (5)$$
$$z = z_1 + t(z_2 - z_1)$$

Ecuaciones paramétricas de una recta

Las Ecuaciones (5) se conocen como las *ecuaciones paramétricas* de una recta.

Finalmente, despejando t en (5) y definiendo $x_2 - x_1 = a$, $y_2 - y_1 = b$ y $z_2 - z_1 = c$, encontramos que

$$\frac{x - x_1}{a} = \frac{y - y_1}{b} = \frac{z - z_1}{c} \qquad (6)$$

Ecuaciones simétricas de una recta

Las Ecuaciones (6) se conocen como las *ecuaciones simétricas* de la recta. Aqui a, b y c son los números directores del vector $\mathbf{v}$. Desde luego, las Ecuaciones (6) son válidas sólo si a, b y c son distintas de cero.

Ejemplo 1 Encuentre una ecuación vectorial, las ecuaciones paramétricas y las ecuaciones simétricas de la recta L que pasa por los puntos $P = (2, -1, 6)$ y $Q = (3, 1, -2)$.

Solución Primero calculamos: $\mathbf{v} = (3-2)\mathbf{i} + [1-(-1)]\mathbf{j} + (-2-6)\mathbf{k} = \mathbf{i} + 2\mathbf{j} - 8\mathbf{k}$. Entonces, de (4) si $R = (x, y, z)$ está en la línea, obtenemos $\overrightarrow{OR} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k} = \overrightarrow{OP} + t\mathbf{v} = 2\mathbf{i} - \mathbf{j} + 6\mathbf{k} + t(\mathbf{i} + 2\mathbf{j} - 8\mathbf{k})$ o

$$x = 2 + t \qquad y = -1 + 2t \qquad z = 6 - 8t$$

Finalmente, como $a = 1$, $b = 2$ y $c = -8$, encontramos las ecuaciones simétricas

$$\frac{x-2}{1} = \frac{y+1}{2} = \frac{z-6}{-8} \qquad (7)$$

Para verificar esto veamos que $(2, -1, 6)$ y $(3, 1, -2)$ están realmente en la recta. Tenemos [después de sustituir estos puntos en (7)]

$$\frac{2-2}{1} = \frac{-1+1}{2} = \frac{6-6}{-8} = 0$$

$$\frac{3-2}{1} = \frac{1+1}{2} = \frac{-2-6}{-8} = 1$$

Se pueden encontrar otros puntos de la recta. Si $t = 3$, por ejemplo, obtenemos

$$3 = \frac{x-2}{1} = \frac{y+1}{2} = \frac{z-6}{-8}$$

que nos da el punto $(5, 5, -18)$.

Ejemplo 2 Encuentre las ecuaciones simétricas de la recta que pasa por el punto $(1, -2, 4)$ y es paralela al vector $\mathbf{v} = \mathbf{i} + \mathbf{j} - \mathbf{k}$.

Solución Usamos la fórmula (6) con $P = (x_1, y_1, z_1) = (1, -2, 4)$ y $\mathbf{v}$ como antes, de forma que $a = 1$, $b = 1$ y $c = -1$. Esto da

$$\frac{x-1}{1} = \frac{y+2}{1} = \frac{z-4}{-1}$$

¿Qué sucede si es cero uno de los números directores a, b o c?

Ejemplo 3 Encuentre las ecuaciones simétricas de la recta que contiene a los puntos $P = (3, 4, -1)$ y $Q = (-2, 4, 6)$.

Solución Aquí $\mathbf{v} = -5\mathbf{i} + 7\mathbf{k}$ y $a = -5$, $b = 0$ y $c = 7$. Entonces una representación paramétrica de la línea es $x = 3 - 5t$, $y = 4$ y $z = -1 + 7t$. Despejando t, encontramos que

$$\frac{x-3}{-5} = \frac{z+1}{7} \quad \text{y} \quad y = 4$$

La ecuación $y = 4$ es la ecuación de un plano paralelo al plano xz y así hemos obtenido una ecuación de una recta en ese plano.

Ejemplo 4 Encuentre las ecuaciones simétricas de la recta (en el plano xy) que pasa por los puntos $(x_1, y_1, 0)$ y $(x_2, y_2, 0)$, donde $x_1 \neq x_2$.

Solución En este caso, $\mathbf{v} = (x_2 - x_1)\mathbf{i} + (y_2 - y_1)\mathbf{j}$ y obtenemos

$$\frac{x-x_1}{x_2-x_1}=\frac{y-y_1}{y_2-y_1} \quad \text{y} \quad z=0$$

Podemos escribir esto como

$$y-y_1=\left(\frac{y_2-y_1}{x_2-x_1}\right)(x-x_1)$$

Aquí $(y_2 - y_1)/(x_2 - x_1) = m$, la pendiente de la recta. Cuando $x = 0$, $y = y_1 - [(y_2 - y_1)/(x_2 - x_1)]x_1 = b$, la intersección de la recta con el eje y. Esto es, $y = mx + b$, que es la forma simplificada de una recta en el plano xy. Así vemos que las ecuaciones simétricas de una recta en el espacio realmente son una generalización de la ecuación de una recta en el plano.

¿Qué sucede si dos de los números directores son cero?

Ejemplo 5 Encuentre las ecuaciones simétricas de la recta que pasa por los puntos $P = (2, 3, -2)$ y $Q = (2, -1, -2)$.

Solución Aquí $\mathbf{v} = -4\mathbf{j}$, de manera que $a=0$, $b=-4$ y $c=0$. Una representación paramétrica de la recta es, por la Ecuación (5), $x=2$, $y=3-4t$, $z=-2$. Sabemos que $x=2$ es la ecuación de un plano paralelo al plano yz mientras que $z=-2$ es la ecuación de un plano paralelo al plano xy. Su intersección es la línea $x=2$, $z=-2$, que es paralela al eje y. De hecho, la ecuación $y=3-4t$ dice, esencialmente, que y puede tomar cualquier valor (mientras x y z permanecen fijas).

Advertencia. Las ecuaciones paramétricas o simétricas de una recta *no* son únicas. Para ver esto basta empezar con otros dos puntos de la recta.

Ejemplo 6 En el Ejemplo 1 la recta contenía al punto $(5, 5, -18)$. Escojamos ahora $P = (5, 5, -18)$ y $Q = (3, 1, -2)$. Encontramos que $\mathbf{v} = -2\mathbf{i} - 4\mathbf{j} + 16\mathbf{k}$, así que $x = 5 - 2t$, $y = 5 - 4t$ y $z = -18 + 16t$. (Notemos que si $t = \frac{3}{2}$, entonces $(x, y, z) = (2, -1, 6)$.) Las ecuaciones simétricas ahora son

$$\frac{x-5}{-2}=\frac{y-5}{-4}=\frac{z+18}{16}$$

La ecuación de una recta en el espacio se obtiene especificando un punto en la recta y un vector *paralelo* a esta recta. Podemos deducir la ecuación de un plano en el espacio especificando un punto en el plano y un vector ortogonal a cada vector del plano. Este vector ortogonal se conoce como *vector normal* y se denota por $\mathbf{n}$ (Figura 3.34).

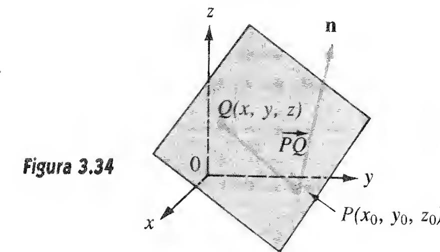


Figura 3.34

Definición 1 **Plano.** Sea P un punto en el espacio y sea $\mathbf{n}$ un vector dado distinto de cero. Entonces el conjunto de todos los puntos Q para los cuales $\overrightarrow{PQ} \cdot \mathbf{n} = 0$ forman un *plano* en $\mathbb{R}^3$.

Notación Usualmente denotamos un plano por π.

Sea $P=(x_0, y_0, z_0)$ un punto fijo en un plano con vector normal $\mathbf{n}=a\mathbf{i}+b\mathbf{j}+c\mathbf{k}$. Si $Q=(x, y, z)$ es cualquier otro punto en el plano entonces $\overrightarrow{PQ}=(x - x_0)\mathbf{i} + (y - y_0)\mathbf{j} + (z - z_0)\mathbf{k}$. Como $\overrightarrow{PQ} \perp \mathbf{n}$, tenemos que $\overrightarrow{PQ} \cdot \mathbf{n} = 0$. Pero esto implica que

$$a(x-x_0)+b(y-y_0)+c(z-z_0)=0 \tag{8}$$

Una forma más común de escribir la ecuación de un plano se deriva fácilmente de (8):

$$ax+by+cz=d \tag{9}$$

en donde

$$d=ax_0+by_0+cz_0=\overrightarrow{OP}\cdot\mathbf{n}$$

Ejemplo 7 Encuentre el plano π que pasa por el punto $(2, 5, 1)$ y tiene por vector normal al $\mathbf{n} = \mathbf{i} - 2\mathbf{j} + 3\mathbf{k}$.

Solución De (8) obtenemos inmediatamente $(x-2)-2(y-5)+3(z-1)=0$ o bien

$$x-2y+3z=-5 \tag{10}$$

Este plano está representado en la Figura 3.35.

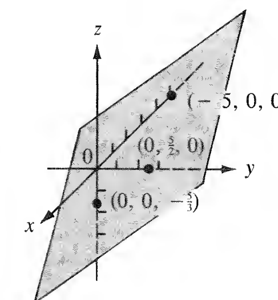


Figura 3.35

Observación. El plano se dibuja fácilmente haciendo $x = y = 0$ en la Ecuación (10) para obtener $(0, 0, -\frac{5}{3})$; $x = z = 0$ para obtener $(0, \frac{5}{2}, 0)$; $y = z = 0$ para obtener $(-5, 0, 0)$. Estos tres puntos están en el plano.

Los tres planos coordenados se representan como sigue:

i. El *plano xy* pasa por el origen (0, 0, 0) y cualquier vector en el eje z es normal a él. De tales vectores el más simple es **k**. Así, de (8), obtenemos $0(x-0)+0(y-0)+1(z-0)=0$, lo que nos lleva a

$$z=0 \tag{11}$$

como la ecuación del plano xy. (Este resultado no debería ser muy sorprendente.)

ii. El *plano xz* tiene la ecuación

$$y=0 \tag{12}$$

iii. El *plano yz* tiene la ecuación

$$x=0 \tag{13}$$

Tres puntos no colineales determinan un plano pues ellos determinan dos vectores no paralelos que se intersecan en un punto (Figura 3.36).

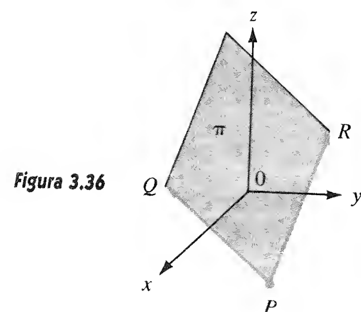


Figura 3.36

Ejemplo 8 Encuentre la ecuación del plano que pasa por los puntos $P = (1, 2, 1)$, $Q = (-2, 3, -1)$ y $R = (1, 0, 4)$.

Solución Los vectores $\overrightarrow{PQ} = -3\mathbf{i}+\mathbf{j}-2\mathbf{k}$ y $\overrightarrow{QR} = 3\mathbf{i}-3\mathbf{j}+5\mathbf{k}$ están en el plano y por tanto son ortogonales al vector normal, de donde

$$\mathbf{n} = \overrightarrow{PQ}\times\overrightarrow{QR} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -3 & 1 & -2 \\ 3 & -3 & 5 \end{vmatrix} = -\mathbf{i}+9\mathbf{j}+6\mathbf{k}$$

y obtenemos

$$\pi: \quad -(x-1)+9(y-2)+6(z-1)=0$$

o

$$-x+9y+6z=23$$

Notemos que si escogemos otro punto, digamos Q, obtenemos la ecuación $-(x+2)+9(y-3)+6(z+1)=0$, que se reduce a $-x+9y+6z=23$. El plano se dibuja en la Figura 3.37.

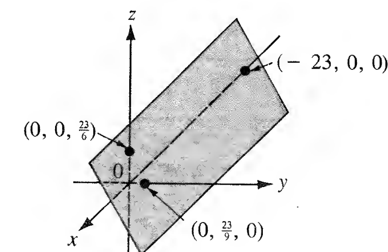


Figura 3.37

Definición 2 **Planos paralelos.** Dos planos son *paralelos** si sus vectores normales son paralelos; esto es, si el producto cruz de sus vectores normales es cero.

Dos planos paralelos se dibujan en la Figura 3.38.

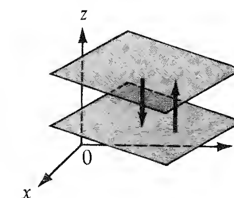


Figura 3.38

Ejemplo 9 Los planos π_1: $2x+3y-z=3$ y π_2: $-4x-6y+2z=8$ son paralelos pues $\mathbf{n}_1 = 2\mathbf{i}+3\mathbf{j}-\mathbf{k}$, $\mathbf{n}_2 = -4\mathbf{i}-6\mathbf{j}+2\mathbf{k} = -2\mathbf{n}_1$ (y $\mathbf{n}_1 \times \mathbf{n}_2 = 0$).

Si dos planos no son paralelos entonces se intersecan en una línea recta.

Ejemplo 10 Encuentre todos los puntos de intersección de los planos $2x-y-z=3$ y $x+2y+3z=7$.

Solución Cuando los planos se intersecan tenemos $x+2y+3z=7$ y $2x-y-z=3$. Resolviendo este sistema de dos ecuaciones en tres incógnitas por reducción por renglón obtenemos, sucesivamente,

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 7 \\ 2 & -1 & -1 & 3 \end{array}\right) \xrightarrow{A_{1,2}(-2)} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 7 \\ 0 & -5 & -7 & -11 \end{array}\right)$$

$$\xrightarrow{M_2(-\frac{1}{5})} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 7 \\ 0 & 1 & \frac{7}{5} & \frac{11}{5} \end{array}\right) \xrightarrow{A_{2,1}(-2)} \left(\begin{array}{ccc|c} 1 & 0 & \frac{1}{5} & \frac{13}{5} \\ 0 & 1 & \frac{7}{5} & \frac{11}{5} \end{array}\right)$$

* Notemos que dos planos paralelos podrían ser coincidentes. Por ejemplo, los planos $x+y+z=1$ y $2x+2y+2z=2$ son coincidentes (el mismo).

Así $y = \frac{11}{5} - (\frac{7}{5})z$ y $x = \frac{13}{5} - (\frac{1}{5})z$. Finalmente, haciendo $z = t$, obtenemos la representación paramétrica de la recta de intersección: $x = \frac{13}{5} - \frac{1}{5}t$, $y = \frac{11}{5} - \frac{7}{5}t$ y $z = t$.

Problemas 3.5

En los Problemas del 1 *al* 14 *encuentre una ecuación vectorial, las ecuaciones paramétricas y las ecuaciones simétricas de la recta indicada.*

1. Que contenga a los puntos (2, 1, 3) y (1, 2, −1)
2. Que contenga a los puntos (1, −1, 1) y (−1, 1, −1)
3. Que contenga a los puntos (−4, 1, 3) y (−4, 0, 1)
4. Que contenga a los puntos (2, 3, −4) y (2, 0, −4)
5. Que contenga a los puntos (1, 2, 3) y (3, 2, 1)
6. Que contenga a los puntos (7, 1, 3) y (−1, −2, 3)
7. Que contenga al punto (2, 2, 1) y paralela a $2\mathbf{i}-\mathbf{j}-\mathbf{k}$
8. Que contenga al punto (−1, −6, 2) y paralela a $4\mathbf{i}+\mathbf{j}-3\mathbf{k}$
9. Que contenga al punto (−1, −2, 5) y paralela a $-3\mathbf{j}+7\mathbf{k}$
10. Que contenga al punto (−2, 3, −2) y paralela a $4\mathbf{k}$
11. Que contenga al punto (a, b, c) y paralela a $d\mathbf{i} + e\mathbf{j}$
12. Que contenga al punto (a, b, c) y paralela a $d\,\mathbf{k}$
13. Que contenga al punto (4, 1, −6) y paralela a $(x-2)/3=(y+1)/6=(z-5)/2$
14. Que contenga al punto (3, 1, −2) y paralela a $(x+1)/3=(y+3)/2=(z-2)/(-4)$
15. Sea L_1 dada por

$$\frac{x-x_1}{a_1}=\frac{y-y_1}{b_1}=\frac{z-z_1}{c_1}$$

y L_2 dada por

$$\frac{x-x_1}{a_2}=\frac{y-y_1}{b_2}=\frac{z-z_1}{c_2}$$

Muestre que L_1 es ortogonal a L_2 si y sólo si $a_1a_2+b_1b_2+c_1c_2=0$.

16. Demuestre que las rectas

$$L_1:\ \frac{x-3}{2}=\frac{y+1}{4}=\frac{z-2}{-1} \quad \text{y} \quad L_2:\ \frac{x-3}{5}=\frac{y+1}{-2}=\frac{z-3}{2}$$

son ortogonales.

17. Demuestre que las rectas

$$L_1:\ \frac{x-1}{1}=\frac{y+3}{2}=\frac{z+3}{3} \quad \text{y} \quad L_2:\ \frac{x-3}{3}=\frac{y-1}{6}=\frac{z-8}{9}$$

son paralelas.

Las rectas en $\mathbb{R}^3$ *que no tienen la misma dirección no necesariamente tienen un punto en común.*

18. Muestre que las rectas L_1: $x = 1 + t$, $y = -3 + 2t$, $z = -2 - t$ y L_2: $x = 17 + 3s$, $y = 4 + s$, $z = -8 - s$ tienen en común al punto (2, −1, −3).
19. Pruebe que las rectas L_1: $x = 2 - t$, $y = 1 + t$, $z = -2t$ y L_2: $x = 1 + s$, $y = -2s$, $z = 3 + 2s$ *no tienen* ningún punto en común.

20. Sea L dada en su forma vectorial $\overrightarrow{OR} = \overrightarrow{OP} + t\mathbf{v}$. Encuentre un número t tal que $\overrightarrow{OR}$ sea perpendicular a $\mathbf{v}$.
21. Utilice el resultado del Problema 20 para encontrar la distancia de la recta L (que contiene a P y es paralela a $\mathbf{v}$) al origen cuando:
    a. $P=(2, 1, -4)$; $\mathbf{v}=\mathbf{i}+\mathbf{j}+\mathbf{k}$
    b. $P=(1, 2, -3)$; $\mathbf{v}=3\mathbf{i}-\mathbf{j}-\mathbf{k}$
    c. $P=(-1, 4, 2)$; $\mathbf{v}=-\mathbf{i}+\mathbf{j}+2\mathbf{k}$

En los Problemas del 22 *al* 25 *encuentre una recta L ortogonal a las dos rectas dadas y que pase a través del punto dado.*

22. $\frac{x+2}{-3}=\frac{y-1}{4}=\frac{z}{-5}$; $\frac{x-3}{7}=\frac{y+2}{-2}=\frac{z-8}{3}$; (1, −3, 2)
23. $\frac{x-2}{-4}=\frac{y+3}{-7}=\frac{z+1}{3}$; $\frac{x+2}{3}=\frac{y-5}{-4}=\frac{z+3}{-2}$; (−4, 7, 3)
24. $x=3-2t$; $y=4+3t$; $z=-7+5t$; $x=-2+4s$, $y=3-2s$, $z=3+s$; (−2, 3, 4)
25. $x=4+10t$, $y=-4-8t$, $z=3+7t$; $x=-2t$, $y=1+4t$, $z=-7-3t$; (4, 6, 0)

★ 26. Calcule la distancia entre las rectas

$$L_1:\ \frac{x-2}{3}=\frac{y-5}{2}=\frac{z-1}{-1} \quad \text{y} \quad L_2:\ \frac{x-4}{-4}=\frac{y-5}{4}=\frac{z+2}{1}$$

[*Sugerencia:* La distancia se mide a lo largo de un vector $\mathbf{v}$ que es perpendicular tanto a L_1 como a L_2. Sean P un punto en L_1 y Q un punto en L_2. Entonces la magnitud de la proyección de $\overrightarrow{PQ}$ en $\mathbf{v}$ es la distancia entre las rectas medida a lo largo de un vector perpendicular a ambas.]

★ 27. Encuentre la distancia entre las rectas

$$L_1:\ \frac{x+2}{3}=\frac{y-7}{-4}=\frac{z-2}{4} \quad \text{y} \quad L_2:\ \frac{x-1}{-3}=\frac{y+2}{4}=\frac{z+1}{1}$$

En los Problemas del 28 *al* 41 *encuentre las ecuaciones del plano.*

28. $P=(0, 0, 0)$; $\mathbf{n}=\mathbf{i}$
29. $P=(0, 0, 0)$; $\mathbf{n}=\mathbf{j}$
30. $P=(0, 0, 0)$; $\mathbf{n}=\mathbf{k}$
31. $P=(1, 2, 3)$; $\mathbf{n}=\mathbf{i}+\mathbf{j}$
32. $P=(1, 2, 3)$; $\mathbf{n}=\mathbf{i}+\mathbf{k}$
33. $P=(1, 2, 3)$; $\mathbf{n}=\mathbf{j}+\mathbf{k}$
34. $P=(2, -1, 6)$; $\mathbf{n}=3\mathbf{i}-\mathbf{j}+2\mathbf{k}$
35. $P=(-4, -7, 5)$; $\mathbf{n}=-3\mathbf{i}-4\mathbf{j}+\mathbf{k}$
36. $P=(-3, 11, 2)$; $\mathbf{n}=4\mathbf{i}+\mathbf{j}-7\mathbf{k}$
37. $P=(3, -2, 5)$; $\mathbf{n}=2\mathbf{i}-7\mathbf{j}-8\mathbf{k}$
38. Que contenga a los puntos (1, 2, −4), (2, 3, 7) y (4, −1, 3)
39. Que contenga a los puntos (−7, 1, 0), (2, −1, 3) y (4, 1, 6)
40. Que contenga a los puntos (1, 0, 0), (0, 1, 0) y (0, 0, 1)
41. Que contenga a los puntos (2, 3, −2), (4, −1, −1) y (3, 1, 2)

Dos planos son ortogonales *si sus vectores normales lo son también. En los Problemas del* 42 *al* 46 *diga si los planos dados son paralelos, ortogonales, coincidentes (esto es, el mismo) o si no se cumple ninguno de estos casos.*

42. π_1: $x+y+z=2$; π_2: $2x+2y+2z=4$
43. π_1: $x-y+z=3$; π_2: $-3x+3y-3z=-9$
44. π_1: $2x-y+z=3$; π_2: $x+y-z=7$
45. π_1: $2x-y+z=3$; π_2: $x+y+z=3$
46. π_1: $3x-2y+7z=4$; π_2: $-2x+4y+2z=16$

En los Problemas del 47 *al* 49 *encuentre la ecuación del conjunto de todos los puntos de intersección de los dos planos.*

47. π_1: $x-y+z=2$; π_2: $2x-3y+4z=7$
48. π_1: $3x-y+4z=3$; π_2: $-4x-2y+7z=8$
★ 49. π_1: $-2x-y+17z=4$; π_2: $2x-y-z=-7$

50. Sea π un plano, P un punto del mismo, n un vector normal al plano y Q un punto fuera del plano (ver la Figura 3.39). Demuestre que la distancia (perpendicular) D de Q al plano está dada por la fórmula

$$D = |\text{proy}_{\mathbf{n}}\,\overrightarrow{PQ}| = \frac{|\overrightarrow{PQ}\cdot \mathbf{n}|}{|\mathbf{n}|}$$

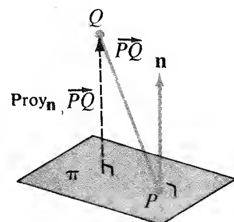


Figura 3.39

En los Problemas del 51 *al* 53 *encontrar la distancia del punto dado del plano dado.*

51. (4, 0, 1); $2x - y + 8z = 3$
52. $(-7, -2, -1)$; $-2x + 8z = -5$
53. $(-3, 0, 2)$; $-3x + y + 5z = 0$

54. Demuestre que la distancia entre el plano $ax + by + cz = d$ y el punto (x_0, y_0, z_0) está dada por

$$D = \frac{|ax_0 + by_0 + cz_0 - d|}{\sqrt{a^2+b^2+c^2}}$$

El ángulo entre dos planos *se define como el ángulo agudo* entre sus vectores normales. En los Problemas del* 55 *al* 57 *halle el valor del ángulo entre los dos planos.*

55. Los planos del Problema 47.
56. Los planos del Problema 48.
57. Los planos del Problema 49.

★ 58. Sean **u** y **v** dos vectores distintos de cero, no paralelos, en un plano π. Muestre que si **w** es otro vector cualquiera en π entonces existen escalares α y β tales que $\mathbf{w} = \alpha\mathbf{u} + \beta\mathbf{v}$. Ésta se conoce como la *representación paramétrica* del plano π. [*Sugerencia:* Dibuje un paralelogramo en el cual $\alpha\mathbf{u}$ y $\beta\mathbf{v}$ formen los lados adyacentes y el vector diagonal sea **w**.]

★ 59. Se dice que tres vectores **u**, **v** y **w** son *coplanares* si los tres están en el mismo plano π. Muestre que si **u**, **v** y **w** pasan por el origen entonces son coplanares si y sólo si el triple producto escalar $\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w})$ es cero.

* Recuerde que un ángulo agudo α es un ángulo entre 0° y 90°, esto es, $\alpha \in [0, \pi/2)$.

En los Problemas del 60 *al* 64 *diga si los tres vectores de posición dados (esto es, con uno de sus puntos finales en el origen) son coplanares. Si son coplanares encuentre la ecuación del plano que los contiene.*

60. $\mathbf{u}=2\mathbf{i}-3\mathbf{j}+4\mathbf{k}$; $\mathbf{v}=7\mathbf{i}-2\mathbf{j}+3\mathbf{k}$; $\mathbf{w}=9\mathbf{i}-5\mathbf{j}+7\mathbf{k}$
61. $\mathbf{u}=-3\mathbf{i}+\mathbf{j}+8\mathbf{k}$; $\mathbf{v}=-2\mathbf{i}-3\mathbf{j}+5\mathbf{k}$; $\mathbf{w}=2\mathbf{i}+14\mathbf{j}-4\mathbf{k}$
62. $\mathbf{u}=2\mathbf{i}+\mathbf{j}-2\mathbf{k}$; $\mathbf{v}=2\mathbf{i}-\mathbf{j}-2\mathbf{k}$; $\mathbf{w}=2\mathbf{i}-\mathbf{j}+2\mathbf{k}$
63. $\mathbf{u}=3\mathbf{i}-2\mathbf{j}+\mathbf{k}$; $\mathbf{v}=\mathbf{i}+\mathbf{j}-5\mathbf{k}$; $\mathbf{w}=-\mathbf{i}+5\mathbf{j}-16\mathbf{k}$
64. $\mathbf{u}=2\mathbf{i}-\mathbf{j}-\mathbf{k}$; $\mathbf{v}=4\mathbf{i}+3\mathbf{j}+2\mathbf{k}$; $\mathbf{w}=6\mathbf{i}+7\mathbf{j}+5\mathbf{k}$

Ejercicios de repaso • Capítulo 3

En los Ejercicios del 1 *al* 6 *encuentre la magnitud y la dirección del vector dado.*

1. $\mathbf{v}=(3, 3)$
2. $\mathbf{v}=-3\mathbf{i}+3\mathbf{j}$
3. $\mathbf{v}=(2, -2\sqrt{3})$
4. $\mathbf{v}=(\sqrt{3}, 1)$
5. $\mathbf{v}=-12\mathbf{i}-12\mathbf{j}$
6. $\mathbf{v}=\mathbf{i}+4\mathbf{j}$

En los Ejercicios del 7 *al* 10 *escriba el vector representado por* $\overrightarrow{PQ}$ *en la forma* $a\mathbf{i} + b\mathbf{j}$. *Dibuje* $\overrightarrow{PQ}$ *y* **v**.

7. $P=(2, 3)$; $Q=(4, 5)$
8. $P=(1, -2)$; $Q=(7, 12)$
9. $P=(-1, -6)$; $Q=(3, -4)$
10. $P=(-1, 3)$; $Q=(3, -1)$
11. Sean $\mathbf{u} = (2, 1)$ y $\mathbf{v} = (-3, 4)$. Determine: (a) $5\mathbf{u}$; (b) $\mathbf{u} - \mathbf{v}$; (c) $-8\mathbf{u} + 5\mathbf{v}$.
12. Sean $\mathbf{u} = -4\mathbf{i} + \mathbf{j}$ y $\mathbf{v} = -3\mathbf{i} - 4\mathbf{j}$. Determine: (a) $-3\mathbf{v}$; (b) $\mathbf{u} + \mathbf{v}$; (c) $3\mathbf{u} - 6\mathbf{v}$.

En los Ejercicios del 13 *al* 19 *encuentre un vector unitario que tenga la misma dirección que el vector dado.*

13. $\mathbf{v}=\mathbf{i}+\mathbf{j}$
14. $\mathbf{v}=-\mathbf{i}+\mathbf{j}$
15. $\mathbf{v}=2\mathbf{i}+5\mathbf{j}$
16. $\mathbf{v}=-7\mathbf{i}+3\mathbf{j}$
17. $\mathbf{v}=3\mathbf{i}+4\mathbf{j}$
18. $\mathbf{v}=-2\mathbf{i}-2\mathbf{j}$
19. $\mathbf{v}=a\mathbf{i}-a\mathbf{j}$
20. Si $\mathbf{v}=4\mathbf{i}-7\mathbf{j}$, encuentre sen θ y cos θ, donde θ es la dirección de **v**.
21. Encuentre un vector unitario con dirección opuesta a la de $\mathbf{v}=5\mathbf{i}+2\mathbf{j}$.
22. Encuentre dos vectores unitarios ortogonales a $\mathbf{v}=\mathbf{i}-\mathbf{j}$.
23. Encuentre un vector unitario con dirección opuesta a la de $\mathbf{v}=10\mathbf{i}-7\mathbf{j}$.

En los Ejercicios del 24 *al* 27 *encuentre un vector* **v** *con la dirección y magnitud dadas.*

24. $|\mathbf{v}|=2$; $\theta=\pi/3$
25. $|\mathbf{v}|=1$; $\theta=\pi/2$
26. $|\mathbf{v}|=4$; $\theta=\pi$
27. $|\mathbf{v}|=7$; $\theta=5\pi/6$

En los Ejercicios del 28 *al* 31 *calcule el producto escalar de los dos vectores y el coseno del ángulo entre ellos.*

28. $\mathbf{u}=\mathbf{i}-\mathbf{j}$; $\mathbf{v}=\mathbf{i}+2\mathbf{j}$
29. $\mathbf{u}=-4\mathbf{i}$; $\mathbf{v}=11\mathbf{j}$
30. $\mathbf{u}=4\mathbf{i}-7\mathbf{j}$; $\mathbf{v}=5\mathbf{i}+6\mathbf{j}$
31. $\mathbf{u}=-\mathbf{i}-2\mathbf{j}$; $\mathbf{v}=4\mathbf{i}+5\mathbf{j}$

En los Ejercicios del 32 *al* 37 *diga si los vectores dados son ortogonales, paralelos o ninguna de las dos cosas. Luego dibuje cada par.*

32. $\mathbf{u}=2\mathbf{i}-6\mathbf{j}$; $\mathbf{v}=-\mathbf{i}+3\mathbf{j}$
33. $\mathbf{u}=4\mathbf{i}-5\mathbf{j}$; $\mathbf{v}=5\mathbf{i}-4\mathbf{j}$
34. $\mathbf{u}=4\mathbf{i}-5\mathbf{j}$; $\mathbf{v}=-5\mathbf{i}+4\mathbf{j}$
35. $\mathbf{u}=-7\mathbf{i}-7\mathbf{j}$; $\mathbf{v}=\mathbf{i}+\mathbf{j}$

36. $\mathbf{u} = -7\mathbf{i} - 7\mathbf{j}$; $\mathbf{v} = -\mathbf{i} + \mathbf{j}$

37. $\mathbf{u} = -7\mathbf{i} - 7\mathbf{j}$; $\mathbf{v} = -\mathbf{i} - \mathbf{j}$

38. Sean $\mathbf{u} = 2\mathbf{i} + 3\mathbf{j}$ y $\mathbf{v} = 4\mathbf{i} + \alpha\mathbf{j}$. Encuentre α tal que
   a. $\mathbf{u}$ y $\mathbf{v}$ sean ortogonales.
   b. $\mathbf{u}$ y $\mathbf{v}$ sean paralelos.
   c. El ángulo entre $\mathbf{u}$ y $\mathbf{v}$ sea $\pi/4$.
   d. El ángulo entre $\mathbf{u}$ y $\mathbf{v}$ sea $\pi/6$.

En los Ejercicios del 39 *al* 44 *calcule* $\text{proy}_{\mathbf{v}}\mathbf{u}$.

39. $\mathbf{u} = 14\mathbf{i}$; $\mathbf{v} = \mathbf{i} + \mathbf{j}$

40. $\mathbf{u} = 14\mathbf{i}$, $\mathbf{v} = \mathbf{i} - \mathbf{j}$

41. $\mathbf{u} = 3\mathbf{i} - 2\mathbf{j}$; $\mathbf{v} = 3\mathbf{i} + 2\mathbf{j}$

42. $\mathbf{u} = 3\mathbf{i} + 2\mathbf{j}$; $\mathbf{v} = \mathbf{i} - 3\mathbf{j}$

43. $\mathbf{u} = 2\mathbf{i} - 5\mathbf{j}$; $\mathbf{v} = -3\mathbf{i} - 7\mathbf{j}$

44. $\mathbf{u} = 4\mathbf{i} - 5\mathbf{j}$; $\mathbf{v} = -3\mathbf{i} - \mathbf{j}$

45. Sean $P = (3, -2)$, $Q = (4, 7)$, $R = (-1, 3)$ y $S = (2, -1)$. Calcule $\text{proy}_{\overrightarrow{PQ}} \overrightarrow{RS}$ y $\text{proy}_{\overrightarrow{RS}} \overrightarrow{PQ}$.

En los Ejercicios del 46 *al* 48 *encuentre la distancia entre los dos puntos dados.*

46. $(4, -1, 7)$; $(-5, 1, 3)$

47. $(-2, 4, -8)$; $(0, 0, 6)$

48. $(2, -7, 0)$; $(0, 5, -8)$

En los Ejercicios del 49 *al* 51 *encuentre la magnitud y los cosenos directores del vector dado.*

49. $\mathbf{v} = 3\mathbf{j} + 11\mathbf{k}$

50. $\mathbf{v} = \mathbf{i} - 2\mathbf{j} - 3\mathbf{k}$

51. $\mathbf{v} = -4\mathbf{i} + \mathbf{j} + 6\mathbf{k}$

52. Encuentre un vector unitario en la dirección de $\overrightarrow{PQ}$, donde $P = (3, -1, 2)$ y $Q = (-4, 1, 7)$.

53. Encuentre un vector unitario cuya dirección sea la opuesta a la de $\overrightarrow{PQ}$, donde $P = (1, -3, 0)$ y $Q = (-7, 1, -4)$.

En los Ejercicios del 54 *al* 61 *sean* $\mathbf{u} = \mathbf{i} - 2\mathbf{j} + 3\mathbf{k}$, $\mathbf{v} = -3\mathbf{i} + 2\mathbf{j} + 5\mathbf{k}$ y $\mathbf{w} = 2\mathbf{i} - 4\mathbf{j} + \mathbf{k}$. *Calcule:*

54. $\mathbf{u} - \mathbf{v}$

55. $3\mathbf{v} + 5\mathbf{w}$

56. $\text{proy}_{\mathbf{v}}\mathbf{w}$

57. $\text{proj}_{\mathbf{w}}\mathbf{u}$

58. $2\mathbf{u} - 4\mathbf{v} + 7\mathbf{w}$

59. $\mathbf{u} \cdot \mathbf{w} - \mathbf{w} \cdot \mathbf{v}$

60. El ángulo entre $\mathbf{u}$ y $\mathbf{v}$

61. El ángulo entre $\mathbf{v}$ y $\mathbf{w}$

En los Ejercicios del 62 *al* 65 *encuentre el producto cruz* $\mathbf{u} \times \mathbf{v}$.

62. $\mathbf{u} = 3\mathbf{i} - \mathbf{j}$; $\mathbf{v} = 2\mathbf{i} + 4\mathbf{k}$

63. $\mathbf{u} = 7\mathbf{j}$; $\mathbf{v} = \mathbf{i} - \mathbf{k}$

64. $\mathbf{u} = 4\mathbf{i} - \mathbf{j} + 7\mathbf{k}$; $\mathbf{v} = -7\mathbf{i} + \mathbf{j} - 2\mathbf{k}$

65. $\mathbf{u} = -2\mathbf{i} + 3\mathbf{j} - 4\mathbf{k}$; $\mathbf{v} = -3\mathbf{i} + \mathbf{j} - 10\mathbf{k}$

66. Encuentre dos vectores unitarios ortogonales tanto a $\mathbf{u} = \mathbf{i} - \mathbf{j} + 3\mathbf{k}$ como a $\mathbf{v} = -2\mathbf{i} - 3\mathbf{j} + 4\mathbf{k}$.

67. Calcule el área del paralelogramo con vértices adyacentes $(1, 4, -2)$, $(-3, 1, 6)$ y $(1, -2, 3)$.

En los Ejercicios del 68 *al* 71 *encuentre una ecuación vectorial, las ecuaciones paramétricas y las ecuaciones simétricas de la recta dada.*

68. Que contenga a los puntos $(3, -1, 4)$ y $(-1, 6, 2)$

69. Que contenga a los puntos $(-4, 1, 0)$ y $(3, 0, 7)$

70. Que contenga al punto $(3, 1, 2)$ y sea paralela a $3\mathbf{i} - \mathbf{j} - \mathbf{k}$.

71. Que contenga al punto $(1, -2, -3)$ y sea paralela a $(x + 1)/5 = (y - 2)/(-3) = (z - 4)/2$.

72. Muestre que las rectas L_1: $x = 3 - 2t$, $y = 4 + t$, $z = -2 + 7t$ y L_2: $x = -3 + s$, $y = 2 - 4s$, $z = 1 + 6s$ no tienen puntos de intersección.

73. Encuentre la distancia del origen a la recta que pasa por el punto $(3, 1, 5)$ y tiene la dirección del vector $\mathbf{v} = 2\mathbf{i} - \mathbf{j} + \mathbf{k}$.

74. Encuentre la ecuación de la recta que pasa por $(-1, 2, 4)$ y es ortogonal a L_1: $(x - 1)/4 = (y + 6)/3 = z/(-2)$ y L_2: $(x + 3)/5 = (y - 1)/1 = (z + 3)/4$.

En los Ejercicios del 75 *al* 77 *encuentre la ecuación del plano que contenga al punto dado y sea ortogonal al vector normal dado.*

75. $P = (1, 3, -2)$; $\mathbf{n} = \mathbf{i} + \mathbf{k}$

76. $P = (1, -4, 6)$; $\mathbf{n} = 2\mathbf{j} - 3\mathbf{k}$

77. $P = (-4, 1, 6)$; $\mathbf{n} = 2\mathbf{i} - 3\mathbf{j} + 5\mathbf{k}$

78. Encuentre la ecuación del plano que contiene a los puntos $(-2, 4, 1)$, $(3, -7, 5)$ y $(-1, -2, -1)$.

79. Encuentre todos los puntos de intersección de los planos π_1: $-x + y + z = 3$ y π_2: $-4x + 2y - 7z = 5$.

80. Obtenga todos los puntos de intersección de los planos π_1: $-4x + 6y + 8z = 12$ y π_2: $2x - 3y - 4z = 5$.

81. Obtenga todos los puntos de intersección de los planos π_1: $3x - y + 4z = 8$ y π_2: $-3x - y - 11z = 0$.

82. Encuentre la distancia del punto $(1, -2, 3)$ al plano $2x - y - z = 6$.

83. Encuentre el ángulo entre los planos del Ejercicio 79.

84. Muestre que los vectores de posición $\mathbf{u} = \mathbf{i} - 2\mathbf{j} + \mathbf{k}$, $\mathbf{v} = 3\mathbf{i} + 2\mathbf{j} - 3\mathbf{k}$ y $\mathbf{w} = 9\mathbf{i} - 2\mathbf{j} - 3\mathbf{k}$ son coplanares y encuentre la ecuación del plano que los contiene.

CAPÍTULO 4

Espacios vectoriales

4.1 Introducción

Según se vio en el capítulo anterior, los conjuntos $\mathbb{R}^2$ (vectores en el plano) y $\mathbb{R}^3$ (vectores en el espacio) tienen propiedades interesantes. Así, si sumamos dos vectores en $\mathbb{R}^2$ obtenemos otro vector en $\mathbb{R}^2$. Sometidos a la suma, los vectores en $\mathbb{R}^2$ son conmutativos y satisfacen la ley asociativa. Si $\mathbf{x} \in \mathbb{R}^2$, entonces $\mathbf{x}+\mathbf{0}=\mathbf{x}$ y $\mathbf{x}+(-\mathbf{x})=\mathbf{0}$. Además, podemos multiplicar los vectores en $\mathbb{R}^2$ por escalares y establecer varias leyes distributivas. Las mismas propiedades también valen en $\mathbb{R}^3$.

A los conjuntos como $\mathbb{R}^2$ y $\mathbb{R}^3$ se los llama *espacios vectoriales*. Intuitivamente, podemos decir que un espacio vectorial es un conjunto de objetos que cumplen con las reglas descritas en el párrafo anterior.

En este capítulo pasaremos de un mundo concreto, en donde resolvíamos ecuaciones y tratábamos con vectores fácilmente visualizables, a un mundo abstracto de espacios vectoriales arbitrarios. Esto es muy ventajoso, ya que, si establecemos una cierta propiedad sobre espacios vectoriales en general, podemos aplicar dicha propiedad a *todc* espacio vectorial. De otra forma, tendríamos que demostrar dicha propiedad una y otra vez para cada nuevo espacio vectorial que estuviéramos tratando (y de hecho, hay una cantidad innumerable de ellos). Además, como se verá más adelante, los teoremas abstractos que demostraremos no son más difíciles que los que ya hemos visto en este libro.

4.2 Definición y propiedades básicas

Espacio vectorial real. Un *espacio vectorial real V* es un conjunto de objetos llamados *vectores*, junto con dos operaciones, llamadas *suma* y *multiplicación por un escalar* que satisfacen los diez axiomas que se enumeran a continuación.

Notación Si **x** y **y** están en V y si α es un número real, entonces escribiremos $\mathbf{x}+\mathbf{y}$ para la suma de **x** y **y** y $\alpha\mathbf{x}$ para el producto escalar de α y **x**.

Antes de enumerar las propiedades de los vectores en un espacio vectorial, hagamos un par de aclaraciones. Primero, si bien es cierto que es muy útil pensar en $\mathbb{R}^2$ o en $\mathbb{R}^3$ cuando tratamos con algún espacio vectorial, es frecuente encontrarse con espacios vectoriales cuya forma es muy diferente de la de estos espacios tan familiares. (Esto lo veremos muy pronto.) Segundo, la Definición 1 se refiere a un espacio vectorial *real*. La palabra "real" significa que los escalares que usamos son reales. Es muy fácil definir un espacio vectorial *complejo* usando números complejos en lugar de números reales. En este libro se trata básicamente con espacios vectoriales reales, pero no es muy difícil hacer generalizaciones a otros conjuntos de escalares.

Axiomas de un espacio vectorial

i. Si $\mathbf{x} \in V$ y $\mathbf{y} \in V$, entonces $\mathbf{x} + \mathbf{y} \in V$ (es decir, V es cerrado para la suma).
ii. Para todos **x**, **y**, **z** en V, $(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$ (ley asociativa de la suma).
iii. Existe un vector $\mathbf{0} \in V$ tal que para todo $\mathbf{x} \in V$, $\mathbf{x} + \mathbf{0} = \mathbf{0} + \mathbf{x} = \mathbf{x}$ (**0** se conoce como *neutro aditivo*).
iv. Si $\mathbf{x} \in V$, existe un vector $-\mathbf{x}$ en V tal que $\mathbf{x} + (-\mathbf{x}) = \mathbf{0}$ ($-\mathbf{x}$ se conoce como el *inverso aditivo* de **x**).
v. Si **x** y **y** están en V, entonces $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$ (ley conmutativa de la suma de vectores).
vi. Si $\mathbf{x} \in V$, y α es un escalar, entonces $\alpha\mathbf{x} \in V$ (se dice que V es cerrado para la multiplicación escalar).
vii. Si **x** y **y** están en V y si α es un escalar, entonces $\alpha(\mathbf{x} + \mathbf{y}) = \alpha\mathbf{x} + \alpha\mathbf{y}$ (primera ley distributiva).
viii. Si $\mathbf{x} \in V$ y si α y β son escalares, entonces $(\alpha + \beta)\mathbf{x} = \alpha\mathbf{x} + \beta\mathbf{x}$ (segunda ley distributiva).
ix. Si $\mathbf{x} \in V$ y si α y β son escalares, entonces $\alpha(\beta\mathbf{x}) = \alpha\beta\mathbf{x}$ (ley asociativa de la multiplicación por escalar).
x. Para todo vector $\mathbf{x} \in V$, $1\mathbf{x} = \mathbf{x}$ (al escalar 1 se le conoce como *neutro multiplicativo*).

Ejemplo 1 **El espacio $\mathbb{R}^n$.** Sea $V = \mathbb{R}^n = \{(x_1, x_2, \ldots, x_n)\colon x_i \in \mathbb{R} \text{ para } i = 1, 2, \ldots, n\}$.

De la Sección 1.3 (véase Teorema 1.3) podemos ver que V satisface todos los axiomas de un espacio vectorial si tomamos a $\mathbb{R}$ como el conjunto de escalares.

Ejemplo 2 Sea $V=\{0\}$. Esto es, V contiene solamente el número 0. Ya que $0+0=1\cdot 0=0+(0+0)=(0+0)+0=0$, vemos que V es un espacio vectorial, llamado frecuentemente espacio vectorial *trivial*.

Ejemplo 3 Sea $V=\{1\}$. Es decir, V contiene solamente el número 1. En este caso, V *no* es un espacio vectorial, puesto que no cumple con el axioma (*i*). Para ver esto, simplemente notemos que $1+1=2 \notin V$.

Ejemplo 4 Sea $V = \{(x, y)\colon y = mx$, donde m es un número real fijo y x es un número real arbitrario$\}$. Esto es, V consiste en todos los puntos que están sobre la recta $y = mx$ que pasa por el origen y que tiene pendiente m. Supóngase que (x_1, y_1) y (x_2, y_2) están en V. Entonces $y_1 = mx_1$, $y_2 = mx_2$, y

$$(x_1, y_1) + (x_2, y_2) = (x_1, mx_1) + (x_2, mx_2) = (x_1 + x_2, mx_1 + mx_2)$$
$$= (x_1 + x_2, m(x_1 + x_2)) \in V.$$

El axioma (*i*) se satisface. Los axiomas (*ii*), (*iii*) y (*v*) son obvios. Más aún,

$$-(x, mx) = (-x, -mx) = (-x, m(-x)) \in V$$

y

$$(x, mx) + (-x, m(-x)) = (0, 0) = \mathbf{0}$$

de modo que el axioma (*iv*) también es satisfecho. Los demás axiomas se verifican con facilidad y se ve que el conjunto de puntos en el plano que quedan sobre la recta que pasa por el origen constituye un espacio vectorial.

Ejemplo 5 Sea $V = \{(x, y)\colon y = 2x + 1, x \in \mathbb{R}\}$. Esto es, V es el conjunto de puntos sobre la recta $y = 2x + 1$. V *no* es un espacio vectorial porque la cerradura no se cumple, como en el Ejemplo 3. Para ver esto, supongamos que (x_1, y_1) y (x_2, y_2) están en V. Entonces,

$$(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2).$$

Si este último vector estuviese en V, tendríamos que

$$y_1 + y_2 = 2(x_1 + x_2) + 1 = 2x_1 + 2x_2 + 1.$$

Pero $y_1 = 2x_1 + 1$ y $y_2 = 2x_2 + 1$, así que

$$y_1 + y_2 = (2x_1 + 1) + (2x_2 + 1) = 2x_1 + 2x_2 + 2.$$

De aquí se concluye que

$$(x_1 + x_2, y_1 + y_2) \notin V \quad \text{si} \quad (x_1, y_1) \in V \quad \text{y} \quad (x_2, y_2) \in V.$$

Ejemplo 6 Sea $V=\{(x,y,z)\colon ax+by+cz=0\}$. Esto es, V es el conjunto de puntos de $\mathbb{R}^3$ que están sobre el plano que pasa por el origen con vector normal (a,b,c). Supongamos que (x_1, y_1, z_1) y (x_2, y_2, z_2) están en V. Entonces $(x_1, y_1, z_1)+(x_2, y_2, z_2)=(x_1 + x_2, y_1 + y_2, z_1 + z_2)\in V$ porque $a(x_1 + x_2) + b(y_1 + y_2) + c(z_1 + z_2) = (ax_1+by_1+cz_1)+(ax_2+by_2+cz_2)= 0 + 0 = 0$; por lo tanto, se satisface el axioma (*i*). Puede verificarse fácilmente la validez de los otros axio-

mas. De esta manera concluimos que el conjunto de puntos que están sobre un plano en $\mathbb{R}^3$ que pasa por el origen es un espacio vectorial.

Ejemplo 7 Sea $V = P_n$, el conjunto de polinomios con coeficientes reales de grado menor o igual que n. Si $p \in P_n$, entonces

$$p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$$

donde cada a_i es real. La suma $p(x) + q(x)$ se define de la manera obvia: Si $q(x) = b_n x^n + b_{n-1}x^{n-1} + \cdots + b_1 x + b_0$, entonces

$$p(x) + q(x) = (a_n + b_n)x^n + (a_{n-1} + b_{n-1})x^{n-1} + \cdots + (a_1 + b_1)x + (a_0 + b_0)$$

Claramente la suma de dos polinomios de grado menor o igual que n es otro polinomio de grado menor o igual que n, por lo que se satisface el axioma (i). Las propiedades (ii) y de la (v) a la (x) son obvias. Si definimos el polinomio cero por $\mathbf{0} = 0x^n + 0x^{n-1} + \cdots + 0x + 0$, entonces claramente $\mathbf{0} \in P_n$ y se satisface el axioma (iii). Finalmente, haciendo $-p(x) = -a_n x^n - a_{n-1}x^{n-1} - \cdots - a_1 x - a_0$, vemos que vale el axioma (iv), de tal manera que P_n es un espacio vectorial real.

Ejemplo 8 Sea $V = C[0, 1]$ = el conjunto de funciones reales continuas definidas sobre el intervalo [0, 1]. Definimos $(f + g)x = f(x) + g(x)$ y $(\alpha f)(x) = \alpha[f(x)]$. Como la suma de dos funciones continuas es continua, se satisface el axioma (i) y los otros axiomas se verifican fácilmente con $\mathbf{0}$ = la función cero y $(-f)(x) = -f(x)$.

Ejemplo 9 Denotemos con $V = M_{34}$ al conjunto de matrices con componentes reales de tamaño 3×4. Es muy fácil verificar que con la suma usual y multiplicación escalar de matrices, M_{34} es un espacio vectorial con el $\mathbf{0}$ definido con la matriz de tamaño 3×4 cuyas componentes son todas cero. Si $A = (a_{ij})$ está en M_{34}, entonces $-A = (-a_{ij})$ también lo está.

Ejemplo 10 De idéntica manera podemos ver que M_{mn}, el conjunto de matrices reales de tamaño $m \times n$, es un espacio vectorial para cualesquiera dos enteros m y n.

Ejemplo 11 Denotemos con S_3 al conjunto de matrices invertibles de tamaño 3×3. Definamos la "suma" $A + B$ como $A + B = AB$. Si A y B son invertibles, entonces AB es invertible (Teorema 1.8.3) de tal forma que se satisface el axioma (i). El axioma (ii) es simplemente la ley asociativa para la multiplicación matricial (Teorema 1.5.2); los axiomas (iii) y (iv) se satisfacen con $\mathbf{0} = I_3$ y $-A = A^{-1}$. Sin embargo, el axioma (v) no se cumple, puesto que normalmente $AB \neq BA$, por lo que S_3 no es un espacio vectorial.

Ejemplo 12 Sea $V = \{(x, y): y \geq 0\}$. V consiste en los puntos de $\mathbb{R}^2$ que están en la parte superior del plano (los primeros dos cuadrantes). Si $y_1 \geq 0$ y $y_2 \geq 0$, entonces $y_1 + y_2 \geq 0$; por lo tanto, si $(x_1, y_1) \in V$ y $(x_2, y_2) \in V$, entonces se tiene $(x_1 + x_2, y_1 + y_2) \in V$. Sin embargo, V no es un espacio vectorial, puesto que el vector (1, 1), por ejemplo, no tiene un inverso en V ya que $(-1, -1) \notin V$. Más aún, el axioma (vi) no se cumple dado que si $(x, y) \in V$, entonces $\alpha(x, y) \notin V$ si $\alpha < 0$.

Ejemplo 13 **El espacio $\mathbb{C}^n$.** Sea $V = \mathbb{C}^n = \{(c_1, c_2, \ldots, c_n): c_i$ es un número complejo para $i = 1, 2, \ldots, n\}$ y el conjunto de escalares es el conjunto de números complejos. Es muy fácil verificar que $\mathbb{C}^n$ es un espacio vectorial.

Según lo sugieren estos ejemplos, existe una gran variedad de espacios vectoriales y muchas clases de conjuntos que *no* son espacios vectoriales. Terminaremos esta sección demostrando algunos resultados elementales sobre espacios vectoriales.

Teorema 1 Sea V un espacio vectorial. Entonces:

i. $\alpha\mathbf{0} = \mathbf{0}$ para todo número real α.
ii. $0 \cdot \mathbf{x} = \mathbf{0}$ para todo $\mathbf{x} \in V$.
iii. Si $\alpha\mathbf{x} = \mathbf{0}$, entonces $\alpha = 0$ o bien $\mathbf{x} = \mathbf{0}$ (o ambos).
iv. $(-1)\mathbf{x} = -\mathbf{x}$ para todo $\mathbf{x} \in V$.

Demostración **i.** Por el axioma (iii), $\mathbf{0} + \mathbf{0} = \mathbf{0}$; y del axioma ($vii$),

$$\alpha(\mathbf{0} + \mathbf{0}) = \alpha\mathbf{0} + \alpha\mathbf{0} = \alpha\mathbf{0} \quad (1)$$

Sumando $-\alpha\mathbf{0}$ a ambos lados de la última ecuación en (1) y usando la ley asociativa (axioma ii), obtenemos

$$[\alpha\mathbf{0} + \alpha\mathbf{0}] + (-\alpha\mathbf{0}) = \alpha\mathbf{0} + (-\alpha\mathbf{0})$$
$$\alpha\mathbf{0} + [\alpha\mathbf{0} + (-\alpha\mathbf{0})] = \mathbf{0}$$
$$\alpha\mathbf{0} + \mathbf{0} = \mathbf{0}$$
$$\alpha\mathbf{0} = \mathbf{0}$$

ii. Esencialmente tenemos el mismo argumento que el usado en la parte (i). Dado que $0 + 0 = 0$, usando el axioma ($viii$) obtenemos que $0\mathbf{x} = (0 + 0)\mathbf{x} = 0\mathbf{x} + 0\mathbf{x}$, o bien $0\mathbf{x} + (-0\mathbf{x}) = 0\mathbf{x} + [0\mathbf{x} + (-0\mathbf{x})]$, o bien $\mathbf{0} = 0\mathbf{x} + \mathbf{0} = 0\mathbf{x}$.
iii. Hagamos $\alpha\mathbf{x} = \mathbf{0}$. Si $\alpha \neq 0$, multiplicando ambos lados de la ecuación por $1/\alpha$, obtenemos $(1/\alpha)(\alpha\mathbf{x}) = (1/\alpha)\mathbf{0} = \mathbf{0}$ (por la parte i). Pero $(1/\alpha)(\alpha\mathbf{x}) = 1\mathbf{x} = \mathbf{x}$ (por el axioma ix), por lo tanto $\mathbf{x} = \mathbf{0}$.
iv. Dado que $1 + (-1) = 0$, usando la parte (ii), obtenemos

$$\mathbf{0} = 0\mathbf{x} = [1 + (-1)]\mathbf{x} = 1\mathbf{x} + (-1)\mathbf{x} = \mathbf{x} + (-1)\mathbf{x} \quad (2)$$

Sumando $-\mathbf{x}$ en ambos lados de (2), llegamos a que

$$\mathbf{0} + (-\mathbf{x}) = \mathbf{x} + (-1)\mathbf{x} + (-\mathbf{x}) = \mathbf{x} + (-\mathbf{x}) + (-1)\mathbf{x}$$
$$= \mathbf{0} + (-1)\mathbf{x} = (-1)\mathbf{x}$$

Por lo tanto, $-\mathbf{x} = (-1)\mathbf{x}$. Notemos que en la ecuación anterior pudimos cambiar el orden de la suma usando la ley conmutativa (axioma *v*). ■

Observación. La parte (*iii*) del Teorema 1 no es tan obvia como parece. Existen ciertos objetos que satisfacen que $xy = 0$ y sin embargo, ni x ni y son cero. Como un ejemplo, tomemos la multiplicación de las matrices 2×2. Si $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ y $B = \begin{pmatrix} 0 & -2 \\ 0 & 0 \end{pmatrix}$, entonces ni A ni B son cero y como puede verificarse fácilmente $AB = 0$, la matriz cero.

Problemas 4.2

En los Problemas del 1 *al* 20 *determine si el subconjunto dado H del espacio vectorial V es un subespacio de V.*

1. El conjunto de matrices diagonales de tamaño $n \times n$ con la suma común de matrices y la multiplicación escalar común.
2. El conjunto de matrices diagonales de tamaño $n \times n$ con la multiplicación (esto es, $A + B = AB$).
3. $\{(x, y): y \leq 0;\, x, y \text{ reales}\}$ con la suma común y la multiplicación escalar de vectores.
4. Los vectores que caen en el primer cuadrante del plano cartesiano.
5. El conjunto de vectores en $\mathbb{R}^3$ de la forma (x, x, x).
6. El conjunto de polinomios de grado 4, con las operaciones del Ejemplo 7.
7. El conjunto de matrices simétricas de $n \times n$ (Sección 1.9) con la suma común y la multiplicación escalar.
8. El conjunto de matrices de 2×2 de la forma $\begin{pmatrix} 0 & a \\ b & 0 \end{pmatrix}$ con la suma común y la multiplicación escalar.
9. El conjunto de matrices de la forma $\begin{pmatrix} 1 & \alpha \\ \beta & 1 \end{pmatrix}$ con las operaciones matriciales de suma y multiplicación escalar.
10. El conjunto formado solamente por el vector (0, 0), con las operaciones comunes de $\mathbb{R}^2$.
11. El conjunto de polinomios de grado $\leq n$ con término constante igual a cero.
12. El conjunto de polinomios de grado $\leq n$ con término constante positivo a_0.
13. El conjunto de funciones continuas en [0, 1] con $f(0) = 0$ y $f(1) = 0$, con las operaciones del Ejemplo 8.
14. El conjunto de puntos en $\mathbb{R}^3$ que están sobre una recta que pasa por el origen.
15. El conjunto de puntos en $\mathbb{R}^3$ que están sobre la recta $x = t + 1$, $y = 2t$, $z = t - 1$.
16. $\mathbb{R}^2$ con la suma definida por $(x_1, y_1) + (x_2, y_2) = (x_1, x_2 + 1, y_1 + y_2 + 1)$ y la multiplicación escalar usual.
17. El conjunto del Problema 16 con la multiplicación escalar definida por $\alpha(x, y) = (\alpha + \alpha x - 1, \alpha + \alpha y - 1)$.
18. El conjunto que consiste en un solo objeto, con la suma definida como *objeto* + *objeto* = *objeto* y multiplicación escalar definida por α(*objeto*) = *objeto*.

*□19. El conjunto de funciones diferenciables definidas sobre [0, 1] con las operaciones del Ejemplo 8.

★ 20. El conjunto de números reales de la forma $a + b\sqrt{2}$, donde a y b son números racionales, con la suma usual de números reales y con la multiplicación escalar definida solamente para escalares racionales.

21. Demuestre que en un espacio vectorial el neutro aditivo es único.
22. Demuestre que en un espacio vectorial cada vector tiene un inverso aditivo único.
23. Si $\mathbf{x}$ y $\mathbf{y}$ son vectores en un espacio vectorial V, demuestre que existe un único vector $\mathbf{z} \in V$, tal que $\mathbf{x} + \mathbf{z} = \mathbf{y}$.
24. Demuestre que el conjunto de números reales positivos forma un espacio vectorial con las operaciones $x + y = xy$ y $\alpha x = x^{\alpha}$.

★ □25. Considere la ecuación diferencial lineal y homogénea de segundo grado

$$y''(x) + a(x)y'(x) + b(x)y(x) = 0$$

donde $a(x)$ y $b(x)$ son funciones continuas. Demuestre que el conjunto de soluciones de la ecuación es un espacio vectorial bajo las reglas usuales de adición de funciones, así como de su multiplicación por números reales.

4.3 Subespacios

En el Ejemplo 4.2.1, vimos que $\mathbb{R}^2 = \{(x, y): x \in \mathbb{R} \text{ y } y \in \mathbb{R}\}$ es un espacio vectorial. En el Ejemplo 4.2.4 vimos que $V = \{(x, y): y = mx\}$ es también un espacio vectorial. Además, es claro que $V \subset \mathbb{R}^2$. Esto es, $\mathbb{R}^2$ tiene un subconjunto que es también un espacio vectorial. De hecho, todos los espacios vectoriales tienen subconjuntos que son a su vez espacios vectoriales. En esta sección estudiaremos estos conjuntos tan importantes.

Definición 1 **Subespacio.** Sea H un subconjunto de un espacio vectorial V y supongamos que H es en sí un espacio vectorial con las operaciones de suma y multiplicación escalar definidas sobre V. Se dice entonces que H es un *subespacio* de V.

En este capítulo veremos muchos ejemplos de subespacios. Antes de ello veremos un resultado mediante el cual resulta relativamente sencillo determinar si un subconjunto de V es de hecho un subespacio de V.

Teorema 1 Un subconjunto no vacío H de un espacio vectorial V es un subespacio de V si las dos reglas de cerradura valen:

Reglas para verificar si un subconjunto es un subespacio.

i. Si $\mathbf{x} \in H$ y $\mathbf{y} \in H$, entonces $\mathbf{x} + \mathbf{y} \in H$.

ii. Si $\mathbf{x} \in H$, entonces $\alpha\mathbf{x} \in H$ para todo escalar α.

* Este símbolo cuadrado se usará en este libro para indicar que el problema requiere métodos de Cálculo.

Demostración Para demostrar que H es un espacio vectorial, debemos verificar que los axiomas (i) a (x) de la página 180 cumplen con las operaciones de la suma vectorial y multiplicación escalar definidas en V. Las dos operaciones de cerradura (axiomas i y vi) se cumplen por hipótesis. Puesto que los vectores en H también están en V, las leyes asociativa, conmutativa, distributiva y la del neutro multiplicativo (axiomas ii, v, vii, $viii$, ix y x) se satisfacen. Ahora, si $\mathbf{x} \in H$ entonces $0\mathbf{x} \in H$, debido a la hipótesis (ii). Pero por el Teorema 4.2.1 (parte ii), $0\mathbf{x} = \mathbf{0}$. Consecuentemente $\mathbf{0} \in H$ y el axioma (iii) se cumple. Finalmente, de la parte (ii) tenemos que $(-1)\mathbf{x} \in H$ para todo $\mathbf{x} \in H$. Por el Teorema 4.2.1 (parte iv), $-\mathbf{x} = (-1)\mathbf{x} \in H$, de tal forma que el axioma (iv) también se cumple y con ello concluimos la demostración. ■

Este teorema nos dice que para probar que H es un subespacio de V, nos basta con verificar que:

> $\mathbf{x}+\mathbf{y}$ y $\alpha\mathbf{x}$ están en H, siempre que $\mathbf{x}$ y $\mathbf{y}$ estén en H y α sea un escalar.

La demostración anterior contiene un resultado importante que debe mencionarse explícitamente:

> Todo subespacio de un espacio vectorial V contiene al $\mathbf{0}$. (1)

Este resultado nos permitirá ver fácilmente si un subes acio particular V *no* es un espacio vectorial. Esto es, si un subconjunto no c ntiene al $\mathbf{0}$, entonces no es un subespacio.

Ahora daremos algunos ejemplos de subespacios.

Ejemplo 1 Para todo espacio vectorial V, el subconjunto $\{\mathbf{0}\}$ que contiene solamente el vector cero, es un subespacio puesto que $\mathbf{0}+\mathbf{0}=\mathbf{0}$ y $\alpha\mathbf{0}=\mathbf{0}$ para todo número real α (parte i del Teorema 4.2.1); se le conoce como el *subespacio trivial.*

Ejemplo 2 V es un subespacio de sí mismo para todo espacio vectorial V.

Subespacio propio Los primeros dos ejemplos nos muestran que todo espacio vectorial V contiene dos espacios vectoriales, $\{\mathbf{0}\}$ y V (a menos que $V=\{\mathbf{0}\}$). Desde luego, es mucho más interesante encontrar otros subespacios. Los subespacios que no son ni $\{\mathbf{0}\}$ ni $\mathbf{V}$ se conocen como *subespacios propios.*

Ejemplo 3 Sea $H = \{(x, y): y = mx\}$ (Ejemplo 4.2.4). Como ya hemos mencionado, H es un subespacio de $\mathbb{R}^2$. Veremos además en la Sección 4.6 (Problema 4.6.15), que los conjuntos de vectores que caen sobre líneas rectas que pasan por el origen son los únicos subespacios propios de $\mathbb{R}^2$.

Ejemplo 4 Sea $H = \{(x, y, z): x = at, y = bt$ y $z = ct$, con a, b, c, t reales$\}$. Esto es, H consiste en los vectores en $\mathbb{R}^3$ que están sobre una recta que pasa por el origen. Veamos que H es un subespacio de $\mathbb{R}^3$. Si $\mathbf{x} = (at_1, bt_1, ct_1) \in H$ y $\mathbf{y} = (at_2, bt_2, ct_2) \in H$, entonces $\mathbf{x} + \mathbf{y} = (a(t_1 + t_2), b(t_1 + t_2), c(t_1 + t_2)) \in H$ y se tiene también que $\alpha\mathbf{x} = (a(\alpha t_1), b(\alpha t_2), c(\alpha t_3)) \in H$. Luego, H es un subespacio de $\mathbb{R}^3$.

Ejemplo 5 Sea $\pi = \{(x, y, z): ax + by + cz = 0; a, b, c$ reales$\}$. Entonces, como vimos en el Ejemplo 4.2.6, π es un espacio vectorial; por lo tanto π es un subespacio de $\mathbb{R}^3$.

En la Sección 4.6 demostraremos que los conjuntos de vectores que están sobre líneas y planos que pasan por el origen son los únicos subespacios propios de $\mathbb{R}^3$.

Antes de estudiar más ejemplos, notemos que *no todo espacio vectorial tiene subespacios propios.*

Ejemplo 6 Sea H un subespacio de $\mathbb{R}$.* Si $H \neq \{\mathbf{0}\}$, entonces H contiene un número real α distinto de cero. Entonces, por el axioma (vi), $1=(1/\alpha)\alpha \in H$ y $\beta 1=\beta \in H$ para todo número real β. Así, si H no es el subespacio trivial, entonces $H=\mathbb{R}$. Esto es, $\mathbb{R}$ *no* tiene subespacios propios.

Ejemplo 7 Si P_n denota el espacio vectorial de polinomios de grado $\leq n$, (Ejemplo 4.2.7), y si $0 < m < n$, entonces P_m es un subespacio propio de P_n, como se puede verificar fácilmente.

Ejemplo 8 Denotemos con M_{mn} (Ejemplo 4.2.10) el espacio vectorial de las matrices de $m \times n$ con componentes reales y sea $H = \{A \in M_{mn}: a_{11} = 0\}$. De las definiciones de suma matricial y de multiplicación escalar, es claro que los dos axiomas de cerradura se satisfacen, de donde H es un subespacio.

Ejemplo 9 Sea $V=M_{nn}$ (las matrices de $n\times n$) y sea $H=\{A \in M_{nn}: A$ es invertible$\}$. Entonces H no es un espacio vectorial puesto que la matriz cero de tamaño $n\times n$ no está en H.

* Note que $\mathbb{R}$ mismo es un espacio vectorial; esto es, $\mathbb{R}$ es un espacio vectorial donde los escalares son los reales. Esto es el Ejemplo 4.2.1 con $n = 1$.

Ejemplo 10 $P_n[0, 1]^* \subset C[0, 1]$ (Ejemplo 4.2.8) puesto que todo polinomio es continuo y P_n es un espacio vectorial para todo entero n, de tal forma que cada $P_n[0,1]$ es un subespacio de $C[0,1]$.

□ **Ejemplo 11** Denotemos con $C'[0,1]$ al conjunto de funciones definidas sobre $[0,1]$ con primera derivada continua. Como toda función derivable es continua, tenemos que $C'[0,1] \subset C[0,1]$. Dado que la suma de dos funciones derivables es derivable y que con la multiplicación por un escalar de una función derivable obtenemos una función derivable, concluimos que $C'[0,1]$ es un subespacio de $C[0,1]$. Además, es un subespacio propio porque no toda función continua es derivable.

□ **Ejemplo 12** Si $f \in C[0,1]$, entonces $\int_0^1 f(x)\,dx$ existe. Sea $H = \{f \in C[0, 1]: \int_0^1 f(x)\,dx = 0\}$. Si $f \in H$ y $g \in H$, entonces $\int_0^1 [f(x)+g(x)]\,dx = \int_0^1 f(x)\,dx + \int_0^1 g(x)\,dx = 0+0=0$ y $\int_0^1 \alpha f(x)\,dx = \alpha\int_0^1 f(x)dx = 0$. Así, $f+g$ y αf están en H para todo número real α. Esto nos muestra que H es un subespacio propio de $C[0,1]$.

Como lo ilustran los tres últimos ejemplos, un espacio vectorial puede tener una gran variedad de subespacios propios. Antes de concluir esta sección, demostraremos un resultado muy interesante sobre subespacios.

Teorema 2 Sean H_1 y H_2 subespacios de un espacio vectorial V. Entonces $H_1 \cap H_2$ es un subespacio de V.

Demostración Sean $\mathbf{x}_1 \in H_1 \cap H_2$ y $\mathbf{x}_2 \in H_1 \cap H_2$. Entonces, dado que H_1 y H_2 son subespacios, $\mathbf{x}_1 + \mathbf{x}_2 \in H_1$ y $\mathbf{x}_1 + \mathbf{x}_2 \in H_2$. Esto es, $\mathbf{x}_1 + \mathbf{x}_2 \in H_1 \cap H_2$. Análogamente, $\alpha\mathbf{x}_1 \in H_1 \cap H_2$. De esta manera se satisfacen los dos axiomas de cerradura y por lo tanto $H_1 \cap H_2$, es un subespacio.† ■

Ejemplo 13 En $\mathbb{R}^3$, sean $H_1 = \{(x, y, z): 2x - y - z = 0\}$ y $H_2 = \{(x, y, z): x + 2y + 3z = 0\}$. Entonces H_1 y H_2 están formados por vectores que están sobre planos que pasan por el origen y, por el Ejemplo 5, son subespacios de $\mathbb{R}^3$. $H_1 \cap H_2$ es la intersección de dos planos y la podemos calcular como en la Sección 3.5:

$$\begin{aligned} x+2y+3z&=0 \\ 2x-\ y-\ z&=0 \end{aligned}$$

o bien, si reducimos por renglones:

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 0 \\ 2 & -1 & -1 & 0 \end{array}\right) \xrightarrow{A_{1,2}(-2)} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 0 \\ 0 & -5 & -7 & 0 \end{array}\right) \xrightarrow{M_2(-\frac{1}{5})}$$

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 0 \\ 0 & 1 & \frac{7}{5} & 0 \end{array}\right) \xrightarrow{A_{2,1}(-2)} \left(\begin{array}{ccc|c} 1 & 0 & \frac{1}{5} & 0 \\ 0 & 1 & \frac{7}{5} & 0 \end{array}\right)$$

* P_n [0, 1] denota el conjunto de polinomios de grado $\le n$ definidos sobre el intervalo [0, 1].
† Note en particular, que como $\mathbf{0} \in H_1$ y $\mathbf{0} \in H_2$, tenemos que $\mathbf{0} \in H_1 \cap H_2$.

En consecuencia, todas las soluciones del sistema homogéneo están dadas por $(-\frac{1}{5}z, -\frac{7}{5}z, z)$. Haciendo $z = t$, obtenemos las ecuaciones paramétricas de la recta L en $\mathbb{R}^3$; $x = -\frac{1}{5}t$, $y = -\frac{7}{5}t$, $z = t$. Y como vimos en el Ejemplo 4, el conjunto de vectores sobre L forma un subespacio de $\mathbb{R}^3$.

Problemas 4.3

En los Problemas del 1 al 20 determine si el subconjunto dado H del espacio vectorial V es un subespacio de V.

1. $V = \mathbb{R}^2$; $H = \{(x, y): y \ge 0\}$
2. $V = \mathbb{R}^2$; $H = \{(x, y): x = y\}$
3. $V = \mathbb{R}^3$; $H =$ el plano xy.
4. $V = \mathbb{R}^2$; $H = \{(x, y): x^2 + y^2 \le 1\}$
5. $V = M_{nn}$; $H = \{D \in M_{nn}: D$ es diagonal$\}$.
6. $V = M_{nn}$; $H = \{T \in M_{nn}: T$ es triangular superior$\}$
7. $V = M_{nn}$; $H = \{S \in M_{nn}: S$ es simétrica$\}$.
8. $V = M_{mn}$; $H = \{A \in M_{mn}: a_{ij} = 0\}$
9. $V = M_{22}$; $H = \left\{A \in M_{22}: A = \begin{pmatrix} a & b \\ -b & c \end{pmatrix}\right\}$
10. $V = M_{22}$; $H = \left\{A \in M_{22}: A = \begin{pmatrix} a & 1+a \\ 0 & 0 \end{pmatrix}\right\}$
11. $V = M_{22}$; $H = \left\{A \in M_{22}: A = \begin{pmatrix} 0 & a \\ b & 0 \end{pmatrix}\right\}$
12. $V = P_4$; $H = \{p \in P_4: \deg p = 4\}$
13. $V = P_4$; $H = \{p \in P_4: p(0) = 0\}$
14. $V = P_n$; $H = \{p \in P_n: p(0) = 0\}$
15. $V = P_n$; $H = \{p \in P_n: p(0) = 1\}$
16. $V = C[0, 1]$; $H = \{f \in C[0, 1]: f(0) = f(1) = 0\}$
17. $V = C[0, 1]$; $H = \{f \in C[0, 1]: f(0) = 2\}$

□ 18. $V = C'[0, 1]$; $H = \{f \in C'[0, 1]: f'(0) = 0\}$

□ 19. $V = C[a, b]$, en donde a y b son números reales y $a < b$; $H = \{f \in C[a, b]: \int_a^b f(x)\,dx = 0\}$

□ 20. $V = C[a, b]$; $H = \{f \in C[a, b]: \int_a^b f(x)\,dx = 1\}$

21. Sean $V = M_{22}$, $H_1 = \{A \in M_{22}: a_{11} = 0\}$ y $H_2 = \left\{A \in M_{22}: A = \begin{pmatrix} -b & a \\ a & b \end{pmatrix}\right\}$.
    a. Demuestre que H_1 y H_2 son subespacios
    b. Describa el subconjunto $H = H_1 \cap H_2$ y pruebe que es un subespacio.

□ 22. Si $V = C[0,1]$, denotemos con H_1 el subespacio del Ejemplo 10 y con H_2 el subespacio del Ejemplo 11. Describa el conjunto $H_1 \cap H_2$ y demuestre que es un subespacio.

23. Si A es una matriz de $n \times m$ y $H = \{\mathbf{x} \in \mathbb{R}^m: A\mathbf{x} = \mathbf{0}\}$. Demuestre que H es un subespacio de $\mathbb{R}^m$. A H se le conoce como el *núcleo* de la matriz A.
24. Hagamos $H = \{\mathbf{x} \in \mathbb{R}^m: A\mathbf{x} \neq \mathbf{0}\}$ en el Problema 23. Demuestre que H no es un subespacio de $\mathbb{R}^m$.
25. Sea $H = \{(x, y, z, w): ax + by + cz + dw = 0\}$, donde a, b, c y d son números reales, no todos nulos. Demuestre que H es un subespacio propio de $\mathbb{R}^4$. A H se le conoce como un *hiperplano* en $\mathbb{R}^4$.
26. Sea $H = \{(x_1, x_2, \ldots, x_n): a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0\}$, donde $a_1, a_2, \ldots, a_n$ son números reales, no todos nulos. Demuestre que H es un subespacio propio de $\mathbb{R}^n$. Al igual que en el Problema 25, a H se le conoce como un *hiperplano* en $\mathbb{R}^n$.
27. Sean H_1 y H_2 dos subespacios de un espacio vectorial V y sea $H_1 + H_2 = \{\mathbf{v}: \mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2$ con $\mathbf{v}_1 \in H_1$ y $\mathbf{v}_2 \in H_2\}$. Demuestre que $H_1 + H_2$ es un subespacio de V.

28. Sean $\mathbf{v}_1$ y $\mathbf{v}_2$ dos vectores en $\mathbb{R}^2$. Demuestre que $H = \{\mathbf{v}: \mathbf{v} = a\mathbf{v}_1 + b\mathbf{v}_2;\ a, b$ números reales$\}$ es un subespacio de $\mathbb{R}^2$.

★ **29.** En el Problema 28, demuestre que si $\mathbf{v}_1$ y $\mathbf{v}_2$ no son colineales, entonces $H = \mathbb{R}^2$.

★ **30.** Sean $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ vectores arbitrarios en un espacio vectorial V. Sea $H = \{\mathbf{v} \in V: \mathbf{v} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_n\mathbf{v}_n$, donde $a_1, a_2, \ldots, a_n$ son escalares$\}$. Demuestre que H es un subespacio de V. A H se le conoce como el subespacio *generado* por los vectores $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$.

4.4 Independencia lineal

En el estudio de álgebra lineal, una de las ideas centrales es la de independencia o dependencia lineal de vectores. En esta sección definiremos lo que entendemos por independencia lineal y se mostrará el modo en que se relaciona con la teoría de sistemas homogéneos de ecuaciones y con los determinantes.

¿Existe una relación especial entre los vectores $\mathbf{v}_1 = \begin{pmatrix}1\\2\end{pmatrix}$ y $\mathbf{v}_2 = \begin{pmatrix}2\\4\end{pmatrix}$? Desde luego, vemos que $\mathbf{v}_2 = 2\mathbf{v}_1$, o escribiendo esto de otro modo, que

$$2\mathbf{v}_1 - \mathbf{v}_2 = \mathbf{0} \tag{1}$$

¿Qué propiedad especial comparten los vectores $\mathbf{v}_1 = \begin{pmatrix}1\\2\\3\end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix}-4\\1\\5\end{pmatrix}$ y $\mathbf{v}_3 = \begin{pmatrix}-5\\8\\19\end{pmatrix}$? Esta pregunta es más difícil de contestar a simple vista. Es fácil verificar que $\mathbf{v}_3 = 3\mathbf{v}_1 + 2\mathbf{v}_2$, o escribiendo de otra manera

$$3\mathbf{v}_1 + 2\mathbf{v}_2 - \mathbf{v}_3 = \mathbf{0} \tag{2}$$

Se ve pues que los dos vectores en la Ecuación (1) y los tres vectores en la Ecuación (2), mantienen una relación más cercana que un par cualquiera de vectores-2 o una terna arbitraria de vectores-3. En cada caso, decimos que los vectores son *linealmente dependientes*. En general, se tiene la siguiente definición importante.

Definición 1 **Dependencia e independencia lineales.** Sean $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ n vectores en un espacio vectorial V. Entonces se dice que los vectores son *linealmente dependientes* si existen n escalares $c_1, c_2, \ldots, c_n$ *no todos cero*, tales que

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n = \mathbf{0} \tag{3}$$

Si los vectores no son linealmente dependientes, entonces se dice que son *linealmente independientes*.

Expresado de otro modo, $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ son linealmente independientes si la ecuación $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n = \mathbf{0}$ sólo se satisface si $c_1 = c_2 = \cdots = c_n = 0$.

¿Cómo se determina si un conjunto de vectores es o no linealmente independiente? El caso de dos vectores es sencillo.

Teorema 1 Dos vectores en un espacio vectorial V son linealmente dependientes si y sólo si uno es un múltiplo del otro.

Demostración Primero suponemos que $\mathbf{v}_2 = c\mathbf{v}_1$ con un escalar $c \neq 0$. Entonces $c\mathbf{v}_1 - \mathbf{v}_2 = \mathbf{0}$; $\mathbf{v}_1$ y $\mathbf{v}_2$ son linealmente dependientes. Por otro lado, si $\mathbf{v}_1$ y $\mathbf{v}_2$ son linealmente dependientes, entonces existen constantes c_1 y c_2, ninguna de las cuales es cero, tales que $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 = \mathbf{0}$. Si $c_1 \neq 0$, entonces, al dividir entre c_1, se obtiene $\mathbf{v}_1 + (c_2/c_1)\mathbf{v}_2 = \mathbf{0}$, o bien

$$\mathbf{v}_1 = \left(-\frac{c_2}{c_1}\right)\mathbf{v}_2$$

Esto es, $\mathbf{v}_1$ es un múltiplo escalar de $\mathbf{v}_2$. Si $c_1 = 0$ entonces $c_2 \neq 0$ y por lo mismo $\mathbf{v}_2 = \mathbf{0} = 0\mathbf{v}_1$. ■

Ejemplo 1 Los vectores $\mathbf{v}_1 = \begin{pmatrix}2\\-1\\0\\3\end{pmatrix}$ y $\mathbf{v}_2 = \begin{pmatrix}-6\\3\\0\\-9\end{pmatrix}$ son linealmente dependientes porque $\mathbf{v}_2 = -3\mathbf{v}_1$.

Ejemplo 2 Los vectores $\begin{pmatrix}1\\2\\4\end{pmatrix}$ y $\begin{pmatrix}2\\5\\-3\end{pmatrix}$ son linealmente independientes; si no lo fueran, tendríamos que $\begin{pmatrix}2\\5\\-3\end{pmatrix} = c\begin{pmatrix}1\\2\\4\end{pmatrix} = \begin{pmatrix}c\\2c\\4c\end{pmatrix}$. Entonces, $2 = c$, $5 = 2c$ y $-3 = 4c$, lo cual es imposible para cualquier número c.

Ejemplo 3 Determine si los vectores $\begin{pmatrix}1\\-2\\3\end{pmatrix}$, $\begin{pmatrix}2\\-2\\0\end{pmatrix}$ y $\begin{pmatrix}0\\1\\7\end{pmatrix}$ son linealmente dependientes o independientes.

Solución Supóngase que $c_1\begin{pmatrix}1\\-2\\3\end{pmatrix} + c_2\begin{pmatrix}2\\-2\\0\end{pmatrix} + c_3\begin{pmatrix}0\\1\\7\end{pmatrix} = \mathbf{0} = \begin{pmatrix}0\\0\\0\end{pmatrix}$. Entonces, multiplicando y sumando, resulta $\begin{pmatrix}c_1 + 2c_2\\ -2c_1 - 2c_2 + c_3\\ 3c_1 + 7c_3\end{pmatrix} = \begin{pmatrix}0\\0\\0\end{pmatrix}$. Esto da un sistema homogéneo de tres ecuaciones con tres incógnitas, c_1, c_2, c_3:

$$\begin{aligned} c_1 + 2c_2 \qquad &= 0\\ -2c_1 - 2c_2 + c_3 &= 0\\ 3c_1 \qquad + 7c_3 &= 0 \end{aligned} \tag{4}$$

Por lo tanto, los vectores serán linealmente dependientes si y sólo si el sistema

(4) posee soluciones no triviales. Se expresa el sistema (4) usando la matriz aumentada y se reduce por filas o renglones:

$$\left(\begin{array}{ccc|c} 1 & 2 & 0 & 0 \\ -2 & -2 & 1 & 0 \\ 3 & 0 & 7 & 0 \end{array}\right) \xrightarrow[A_{1,3}(-3)]{A_{1,2}(2)} \left(\begin{array}{ccc|c} 1 & 2 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & -6 & 7 & 0 \end{array}\right)$$

$$\xrightarrow{M_2(\frac{1}{2})} \left(\begin{array}{ccc|c} 1 & 2 & 0 & 0 \\ 0 & 1 & \frac{1}{2} & 0 \\ 0 & -6 & 7 & 0 \end{array}\right) \xrightarrow[A_{2,3}(6)]{A_{2,1}(-2)} \left(\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & \frac{1}{2} & 0 \\ 0 & 0 & 10 & 0 \end{array}\right)$$

$$\xrightarrow{M_3(\frac{1}{10})} \left(\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & \frac{1}{2} & 0 \\ 0 & 0 & 1 & 0 \end{array}\right) \xrightarrow[A_{3,2}(-\frac{1}{2})]{A_{3,1}(1)} \left(\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right)$$

El último sistema de ecuaciones se lee: $c_1 = 0$, $c_2 = 0$, $c_3 = 0$. De manera que (4) carece de soluciones no triviales y los vectores dados son linealmente independientes.

Ejemplo 4 Determinar si los vectores $\begin{pmatrix} 1 \\ -3 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 0 \\ 4 \end{pmatrix}$ y $\begin{pmatrix} 11 \\ -6 \\ 12 \end{pmatrix}$ son linealmente independientes o no.

Solución La ecuación $c_1\begin{pmatrix} 1 \\ -3 \\ 0 \end{pmatrix} + c_2\begin{pmatrix} 3 \\ 0 \\ 4 \end{pmatrix} + c_3\begin{pmatrix} 11 \\ -6 \\ 12 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ conduce al sistema homogéneo:

$$\begin{aligned} c_1 + 3c_2 + 11c_3 &= 0 \\ -3c_1 \qquad - 6c_3 &= 0 \\ 4c_2 + 12c_3 &= 0 \end{aligned} \tag{5}$$

Escribiendo el sistema (5) en la forma de matriz aumentada y reduciendo por filas, que sucesivamente,

$$\left(\begin{array}{ccc|c} 1 & 3 & 11 & 0 \\ -3 & 0 & -6 & 0 \\ 0 & 4 & 12 & 0 \end{array}\right) \xrightarrow{A_{1,2}(3)} \left(\begin{array}{ccc|c} 1 & 3 & 11 & 0 \\ 0 & 9 & 27 & 0 \\ 0 & 4 & 12 & 0 \end{array}\right)$$

$$\xrightarrow{M_2(\frac{1}{9})} \left(\begin{array}{ccc|c} 1 & 3 & 11 & 0 \\ 0 & 1 & 3 & 0 \\ 0 & 4 & 12 & 0 \end{array}\right) \xrightarrow[A_{2,3}(-4)]{A_{2,1}(-3)} \left(\begin{array}{ccc|c} 1 & 0 & 2 & 0 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Aquí podemos detenernos, pues la teoría de la Sección 1.7 muestra que el sistema (5) posee una infinidad de soluciones. Por ejemplo, la última matriz aumentada da

$$\begin{aligned} c_1 \qquad + 2c_3 &= 0 \\ c_2 + 3c_3 &= 0 \end{aligned}$$

Si elegimos $c_3 = 1$, entonces $c_2 = -3$ y $c_1 = -2$, así que, como se puede verificar fácilmente,

$$-2\begin{pmatrix} 1 \\ -3 \\ 0 \end{pmatrix} - 3\begin{pmatrix} 3 \\ 0 \\ 4 \end{pmatrix} + \begin{pmatrix} 11 \\ -6 \\ 12 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

y los vectores son linealmente dependientes.

La teoría de los sistemas homogéneos expresa algo respecto de la independencia o la dependencia lineales de los vectores.

Teorema 2 Un conjunto de n vectores en $\mathbb{R}^m$ es siempre linealmente dependiente si $n > m$.

Demostración Sean $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ n vectores en $\mathbb{R}^m$ y tratemos de evaluar constantes $c_1, c_2, \ldots, c_n$, no todas nulas, tales que

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n = \mathbf{0} \tag{6}$$

Sean $\mathbf{v}_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}, \mathbf{v}_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix}, \ldots, \mathbf{v}_n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}$. Entonces, la Ecuación (6) se convierte en

$$\begin{aligned} a_{11}c_1 + a_{12}c_2 + \cdots + a_{1n}c_n &= 0 \\ a_{21}c_1 + a_{22}c_2 + \cdots + a_{2n}c_n &= 0 \\ \vdots \qquad \vdots \qquad\quad \vdots \quad &\;\; \vdots \\ a_{m1}c_1 + a_{m2}c_2 + \cdots + a_{mn}c_n &= 0 \end{aligned} \tag{7}$$

Pero el sistema (7) es el sistema (1.7.1) y, de acuerdo con el Teorema 1.7.1, tal sistema tiene una infinidad de soluciones si $n > m$. Así pues, existen escalares $c_1, c_2, \ldots, c_n$ no todos cero que satisfacen (7) y los vectores $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ son linealmente dependientes. ■

Ejemplo 5 Los vectores $\begin{pmatrix} 2 \\ -3 \\ 4 \end{pmatrix}, \begin{pmatrix} 4 \\ 7 \\ -6 \end{pmatrix}, \begin{pmatrix} 18 \\ -11 \\ 4 \end{pmatrix}$ y $\begin{pmatrix} 2 \\ -7 \\ 3 \end{pmatrix}$ son linealmente dependientes, por ser cuatro vectores-3.

Existe un corolario al Teorema 2 muy importante (y obvio).

Corolario Un conjunto linealmente independiente de vectores en $\mathbb{R}^n$ contiene, a lo sumo, n vectores.

Nota. Podemos expresar el corolario como sigue: Si se tienen n vectores-n li-

nealmente independientes, no es posible agregar más vectores sin hacer que el conjunto obtenido sea linealmente dependiente.

El sistema (7) permite formular otra observación importante cuya demostración se deja como ejercicio (vea Problema 27).

Teorema 3 Sea

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

Entonces las columnas de A, consideradas como vectores, son linealmente dependientes si y sólo si el sistema (7), que puede ser expresado como $A\mathbf{c} = \mathbf{0}$, posee un número infinito de soluciones. Aquí, $\mathbf{c} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$.

Ejemplo 6 Considérese el sistema homogéneo

$$\begin{aligned} x_1 + 2x_2 - x_3 + 2x_4 &= 0 \\ 3x_1 + 7x_2 + x_3 + 4x_4 &= 0 \end{aligned} \qquad (8)$$

Resolvamos mediante reducción por renglones:

$$\left(\begin{array}{cccc|c} 1 & 2 & -1 & 2 & 0 \\ 3 & 7 & 1 & 4 & 0 \end{array}\right) \xrightarrow{A_{1,2}(-3)} \left(\begin{array}{cccc|c} 1 & 2 & -1 & 2 & 0 \\ 0 & 1 & 4 & -2 & 0 \end{array}\right)$$

$$\xrightarrow{A_{2,1}(-2)} \left(\begin{array}{cccc|c} 1 & 0 & -9 & 6 & 0 \\ 0 & 1 & 4 & -2 & 0 \end{array}\right)$$

El último sistema es:

$$\begin{aligned} x_1 \qquad - 9x_3 + 6x_4 &= 0 \\ x_2 + 4x_3 - 2x_4 &= 0 \end{aligned}$$

Vemos que este sistema posee un número infinito de soluciones, que escribimos como un vector columna:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 9x_3 - 6x_4 \\ -4x_3 + 2x_4 \\ x_3 \\ x_4 \end{pmatrix} = x_3 \begin{pmatrix} 9 \\ -4 \\ 1 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} -6 \\ 2 \\ 0 \\ 1 \end{pmatrix} \qquad (9)$$

Nótese que $\begin{pmatrix} 9 \\ -4 \\ 1 \\ 0 \end{pmatrix}$ y $\begin{pmatrix} -6 \\ 2 \\ 0 \\ 1 \end{pmatrix}$ son soluciones linealmente independientes de (8) porque ninguno es múltiplo del otro. (El lector debe verificar que los vectores son soluciones.) Como x_3 y x_4 son números reales arbitrarios, resulta de (9) que es posible expresar todas las soluciones de (8) en términos de dos vectores solución linealmente independientes.

Los siguientes dos teoremas se deducen directamente del Teorema 3.

Teorema 4 Sean $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$, n vectores en $\mathbb{R}^n$, y sea A una matriz de $n \times n$ cuyas columnas son $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$. Entonces $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ son linealmente independientes si y sólo si la única solución al sistema homogéneo $A\mathbf{x} = \mathbf{0}$ es la solución trivial $\mathbf{x} = \mathbf{0}$.

Demostración Esto es el Teorema 3 en el caso en que $m = n$. ■

Teorema 5 Sea A una matriz de $n \times n$. Entonces $\det A \neq 0$ si y solamente si las columnas de A son linealmente independientes.

Demostración Del Teorema 4 y del Teorema Resumen (vea pág. 127), las columnas de A son dependientes $\Leftrightarrow \mathbf{0}$ es la única solución de $A\mathbf{x} = \mathbf{0} \Leftrightarrow \det A \neq 0$. (Aquí $\Leftrightarrow$ significa "si y sólo si".) ■

El Teorema 5 permite extender el Teorema Resumen.

Teorema 6 ***Teorema resumen - Versión 5.*** Sea A una matriz de $n \times n$. Entonces cada uno de los siguientes siete enunciados implica a los otros seis (esto es, si uno se verifica, todos son ciertos).

i. A es invertible.
ii. La única solución al sistema homogéneo $A\mathbf{x} = \mathbf{0}$ es la solución trivial $(\mathbf{x} = \mathbf{0})$.
iii. El sistema $A\mathbf{x} = \mathbf{b}$ posee una solución única para todo n-vector $\mathbf{b}$.
iv. A es equivalente por filas a la matriz identidad I_n.
v. A puede ser escrita como el producto de matrices elementales.
vi. $\det A \neq 0$.
vii. Las columnas (y los renglones) de A son linealmente independientes.

Demostración La única parte que no ha sido demostrada es que los renglones de A son linealmente independientes, si $\det A \neq 0$. Si las columnas son independientes, $\det A \neq 0$. Entonces $\det A^t = \det A \neq 0$ (vea Teorema 2.2.3). Así pues, las columnas de A^t son linealmente independientes. Pero las columnas de A^t son los renglones de A, así que los renglones de A son independientes. ■

Todo ejemplo que se ha resuelto hasta ahora ha sido en el espacio $\mathbb{R}^n$. Esto no es tan restrictivo como parece. En la Sección 5.4 (Teorema 6) veremos como muchos espacios vectoriales que parecen distintos, poseen especialmente las mismas propiedades. Por ejemplo, se verá que el espacio P_n es esencialmente el mismo $\mathbb{R}^{n+1}$. Se dice que tales espacios vectoriales son *isomorfos*.

Este poderoso resultado no lo daremos sino hasta el Capítulo 5. Mientras tanto, se examinarán algunos ejemplos en otros espacios que no sean $\mathbb{R}^n$.

Ejemplo 7 En M_{23} sean $A_1=\begin{pmatrix}1&0&2\\3&1&-1\end{pmatrix}$, $A_2=\begin{pmatrix}-1&1&4\\2&3&0\end{pmatrix}$ y $A_3=\begin{pmatrix}-1&0&1\\1&2&1\end{pmatrix}$. Determinar si A_1, A_2, y A_3 son linealmente independientes o no lo son.

Solución Supóngase que $c_1A_1 + c_2A_2 + c_3A_3 = 0$. Entonces

$$\begin{pmatrix}0&0&0\\0&0&0\end{pmatrix}=c_1\begin{pmatrix}1&0&2\\3&1&-1\end{pmatrix}+c_2\begin{pmatrix}-1&1&4\\2&3&0\end{pmatrix}+c_3\begin{pmatrix}-1&0&1\\1&2&1\end{pmatrix}$$

$$=\begin{pmatrix}c_1-c_2-c_3 & c_2 & 2c_1+4c_2+c_3\\3c_1+2c_2+c_3 & c_1+3c_2+2c_3 & -c_1+c_3\end{pmatrix}$$

Esto da un sistema homogéneo de seis ecuaciones en las tres incógnitas c_1, c_2, c_3, y es fácil verificar que la única solución es $c_1 = c_2 = c_3 = 0$. Así pues, las tres matrices son linealmente independientes.

Ejemplo 8 En P_3, determinar si los polinomios $1, x, x^2, x^3$ son linealmente independientes o no.

Solución Supóngase que $c_1 + c_2x + c_3x^2 + c_4x^3 = 0$. Esto debe satisfacerse para cada número real x. En particular, si $x = 0$ obtenemos $c_1 = 0$. Entonces, tomando $x = 1, -1, 2$, obtenemos, sucesivamente:

$$\begin{aligned}c_2+c_3+c_4&=0\\-c_2+c_3-c_4&=0\\2c_2+4c_3+8c_4&=0\end{aligned}$$

El determinante de este sistema homogéneo es

$$\begin{vmatrix}1&1&1\\-1&1&-1\\2&4&8\end{vmatrix}=12\neq 0$$

de modo que el sistema posee la solución única $c_2 = c_3 = c_4 = 0$ y los cuatro polinomios son linealmente independientes. Es posible examinar esto desde otro punto de vista. Sabemos que todo polinomio de grado 3 posee, cuando más, tres raíces reales. Pero si $c_1 + c_2x + c_3x^2 + c_4x^3 = 0$ para algunas constantes no nulas c_1, c_2, c_3 y c_4, y para cada número real x, entonces se ha construido un polinomio cúbico para el cual todo número real es una raíz. Esto es, claramente, imposible.

Ejemplo 9 En P_2, determinar si los polinomios $x - 2x^2$, $x^2 - 4x$, y $-7x + 8x^2$ son linealmente independientes o no.

Solución Sea $c_1(x - 2x^2) + c_2(x^2 - 4x) + c_3(-7x + 8x^2) = 0$. Reordenando términos,

$$(c_1-4c_2-7c_3)x=0$$
$$(-2c_1+c_2+8c_3)x^2=0$$

Estas ecuaciones valen para toda x si y sólo si

$$c_1-4c_2-7c_3=0$$

y

$$-2c_1+c_2+8c_3=0$$

Por el Teorema 1.7.1, este sistema de dos ecuaciones en tres incógnitas posee un número infinito de soluciones. Esto demuestra que los polinomios son linealmente dependientes.

Si resolvemos este sistema homogéneo, se obtiene, sucesivamente:

$$\left(\begin{array}{ccc|c}1&-4&-7&0\\-2&1&8&0\end{array}\right)\xrightarrow{A_{1,2}(2)}\left(\begin{array}{ccc|c}1&-4&-7&0\\0&-7&-6&0\end{array}\right)$$

$$\xrightarrow{M_2(-\frac{1}{7})}\left(\begin{array}{ccc|c}1&-4&-7&0\\0&1&\frac{6}{7}&0\end{array}\right)\xrightarrow{A_{2,1}(4)}\left(\begin{array}{ccc|c}1&0&-\frac{25}{7}&0\\0&1&\frac{6}{7}&0\end{array}\right)$$

Así, c_3 puede ser elegida arbitrariamente, $c_1 = \frac{25}{7}c_3$ y $c_2 = -\frac{6}{7}c_3$. Si $c_3 = 7$, por ejemplo, entonces $c_1 = 25$, $c_2 = -6$, y tendremos $25(x - 2x^2) - 6(x^2 - 4x) + 7(-7x + 8x^2) = 0$.

Problemas 4.4

En los Problemas del 1 al 22 determinar si el conjunto dado de vectores es linealmente independiente o dependiente.

1. $\begin{pmatrix}1\\2\end{pmatrix}$; $\begin{pmatrix}-1\\-3\end{pmatrix}$

2. $\begin{pmatrix}2\\-1\\4\end{pmatrix}$; $\begin{pmatrix}4\\-2\\7\end{pmatrix}$

3. $\begin{pmatrix}2\\-1\\4\end{pmatrix}$; $\begin{pmatrix}4\\-2\\8\end{pmatrix}$

4. $\begin{pmatrix}-2\\3\end{pmatrix}$; $\begin{pmatrix}4\\7\end{pmatrix}$

5. $\begin{pmatrix}-3\\2\end{pmatrix}$; $\begin{pmatrix}1\\10\end{pmatrix}$; $\begin{pmatrix}4\\-5\end{pmatrix}$

6. $\begin{pmatrix}1\\0\\1\end{pmatrix}$; $\begin{pmatrix}0\\1\\1\end{pmatrix}$; $\begin{pmatrix}1\\1\\0\end{pmatrix}$

7. $\begin{pmatrix}1\\0\\0\end{pmatrix}$; $\begin{pmatrix}0\\1\\0\end{pmatrix}$; $\begin{pmatrix}0\\0\\1\end{pmatrix}$

8. $\begin{pmatrix}-3\\4\\2\end{pmatrix}$; $\begin{pmatrix}7\\-1\\3\end{pmatrix}$; $\begin{pmatrix}1\\2\\8\end{pmatrix}$

9. $\begin{pmatrix}-3\\4\\2\end{pmatrix}$; $\begin{pmatrix}7\\-1\\3\end{pmatrix}$; $\begin{pmatrix}1\\1\\8\end{pmatrix}$

10. $\begin{pmatrix}1\\-2\\1\\1\end{pmatrix}$; $\begin{pmatrix}3\\0\\2\\-2\end{pmatrix}$; $\begin{pmatrix}0\\4\\-1\\-1\end{pmatrix}$; $\begin{pmatrix}5\\0\\3\\-1\end{pmatrix}$

11. $\begin{pmatrix}1\\-2\\1\\1\end{pmatrix}; \begin{pmatrix}3\\0\\2\\-2\end{pmatrix}; \begin{pmatrix}0\\4\\-1\\1\end{pmatrix}; \begin{pmatrix}5\\0\\3\\-1\end{pmatrix}$

12. $\begin{pmatrix}1\\-1\\2\end{pmatrix}; \begin{pmatrix}4\\0\\0\end{pmatrix}; \begin{pmatrix}-2\\3\\5\end{pmatrix}; \begin{pmatrix}7\\1\\2\end{pmatrix}.$

13. En P_2: $1-x, x$

14. En P_2: $-x, x^2-2x, 3x+5x^2$

15. En P_2: $1-x, 1+x, x^2$

16. En P_3: x, x^2-x, x^3-x

17. En P_3: $2x, x^3-3, 1+x-4x^3, x^3+18x-9$

18. En M_{22}: $\begin{pmatrix}2&-1\\4&0\end{pmatrix}, \begin{pmatrix}0&-3\\1&5\end{pmatrix}, \begin{pmatrix}4&1\\7&-5\end{pmatrix}$

19. En M_{22}: $\begin{pmatrix}1&-1\\0&6\end{pmatrix}, \begin{pmatrix}-1&0\\3&1\end{pmatrix}, \begin{pmatrix}1&1\\-1&2\end{pmatrix}, \begin{pmatrix}0&1\\1&0\end{pmatrix}$

20. En M_{22}: $\begin{pmatrix}-1&0\\1&2\end{pmatrix}, \begin{pmatrix}2&3\\7&-4\end{pmatrix}, \begin{pmatrix}8&-5\\7&6\end{pmatrix}, \begin{pmatrix}4&-1\\2&3\end{pmatrix}, \begin{pmatrix}2&3\\-1&4\end{pmatrix}$

★ 21. En $C[0, 1]$: sen x, cos x

★ 22. En $C[0, 1]$: $x, \sqrt{x}, \sqrt[3]{x}$

23. Formule una condición para los números a, b, c y d tal que los vectores $\begin{pmatrix}a\\b\end{pmatrix}$ y $\begin{pmatrix}c\\d\end{pmatrix}$ sean linealmente dependientes.

★ 24. Encontrar una condición sobre los números a_{ij} tal que los vectores $\begin{pmatrix}a_{11}\\a_{21}\\a_{31}\end{pmatrix}, \begin{pmatrix}a_{12}\\a_{22}\\a_{32}\end{pmatrix}, \begin{pmatrix}a_{13}\\a_{23}\\a_{33}\end{pmatrix}$ sean linealmente dependientes.

25. ¿Para qué valor(es) de α son linealmente dependientes los vectores $\begin{pmatrix}1\\2\\3\end{pmatrix}, \begin{pmatrix}2\\-1\\4\end{pmatrix}, \begin{pmatrix}3\\\alpha\\4\end{pmatrix}$?

26. ¿Para qué valor(es) de α resultan ser linealmente dependientes los vectores $\begin{pmatrix}2\\-3\\1\end{pmatrix}, \begin{pmatrix}-4\\6\\-2\end{pmatrix}, \begin{pmatrix}\alpha\\1\\2\end{pmatrix}$? [*Sugerencia:* Examíneseles con detenimiento.]

27. Demuestre el Teorema 3. [*Sugerencia:* Examine en detalle el sistema (7).]

28. Demuestre que si los vectores $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ son linealmente dependientes (en $\mathbb{R}^m$), y si $\mathbf{v}_{n+1}$ es cualquier otro vector en $\mathbb{R}^m$, entonces el conjunto $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n, \mathbf{v}_{n+1}$ es linealmente dependiente.

29. Demuestre que si $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ $(n \geq 2)$ son linealmente independientes, entonces también lo son $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$, donde $k < n$.

30. Pruebe que si en $\mathbb{R}^n$ los vectores $\mathbf{v}_1$ y $\mathbf{v}_2$ son ortogonales (vea pág. 00), entonces el conjunto $\{\mathbf{v}_1, \mathbf{v}_2\}$ es linealmente independiente.

★ 31. Supóngase que $\mathbf{v}_1$ es ortogonal respecto de $\mathbf{v}_2$ y $\mathbf{v}_3$, y que $\mathbf{v}_2$ es ortogonal respecto de $\mathbf{v}_3$. Si $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ no son nulos, pruebe que el conjunto $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es linealmente independiente.

32. Sea A una matriz cuadrada $(n \times n)$ con columnas $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$. Demuestre que $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ son linealmente independientes si y sólo si la forma escalón por filas de A no contiene un renglón de ceros.

En los Problemas del 33 al 37 escribir las soluciones a los sistemas homogéneos dados en términos de uno o más vectores linealmente independientes.

33. $x_1 + x_2 + x_3 = 0$

34. $x_1 - x_2 + 7x_3 - x_4 = 0$
$2x_1 + 3x_2 - 8x_3 + x_4 = 0$

35. $x_1 + 2x_2 - x_3 = 0$
$2x_1 + 5x_2 + 4x_3 = 0$

36. $x_1 + x_2 + x_3 - x_4 - x_5 = 0$
$-2x_1 + 3x_2 + x_3 + 4x_4 - 6x_5 = 0$

37. $x_1 + 2x_2 - 3x_3 + 5x_4 = 0$

38. Sea $\mathbf{u} = (1, 2, 3)$.
   a. Sea $H = \{\mathbf{v} \in \mathbb{R}^3: \mathbf{u} \cdot \mathbf{v} = 0\}$. Demuestre que H es un subespacio de $\mathbb{R}^3$.
   b. Halle dos vectores en H linealmente independientes y llámeselos $\mathbf{x}$ y $\mathbf{y}$.
   c. Evaluar $\mathbf{w} = \mathbf{x} \times \mathbf{y}$.
   d. Pruebe que $\mathbf{u}$ y $\mathbf{w}$ son linealmente dependientes.
   e. Dé una interpretación geométrica de las partes (a) y (c) y explique por qué (d) debe verificarse.

Observación. Si $V = \{\mathbf{v} \in \mathbb{R}^3: \mathbf{v} = \alpha\mathbf{u}$ para un número real $\alpha\}$, entonces V es un subespacio de $\mathbb{R}^3$ y H se llama *complemento ortogonal* de V.

39. Elíjase un vector $\mathbf{u} \neq \mathbf{0}$ en $\mathbb{R}^3$. Repítanse los pasos del Problema 38 con el vector elegido.

40. Demuestre que cuatro polinomios en P_2 son linealmente dependientes.

41. Demuestre que dos polinomios no pueden generar todo P_2.

★ 42. Pruebe que $n + 2$ polinomios en P_n son linealmente dependientes.

43. Pruebe que cualquier subconjunto de un conjunto de vectores linealmente independiente es linealmente independiente. [*Nota:* Esto generaliza el Problema 29.]

44. Demuestre que siete matrices en M_{32} son linealmente dependientes.

★ 45. Demuestre que $mn + 1$ matrices en M_{mn} son linealmente dependientes.

46. Sean S_1 y S_2 dos conjuntos de vectores linealmente independientes en un espacio vectorial V. Demuestre que $S_1 \cap S_2$ es linealmente independiente.

★ 47. Demuestre que en P_n los polinomios $1, x, x^2, \ldots, x^n$ son linealmente independientes. [*Sugerencia:* Si $n = 1$, esto se verifica. Supóngase que $1, x, x^2, \ldots, x^{n-1}$ es linealmente independiente y demuestre que esto implica que $1, x, x^2, \ldots, x^n$ también lo es. Lo anterior completa la demostración por inducción matemática.]

48. Sea $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ un conjunto linealmente independiente. Demuestre que los vectores $\mathbf{v}_1, \mathbf{v}_1 + \mathbf{v}_2, \mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}_3, \ldots, \mathbf{v}_1 + \mathbf{v}_2 + \cdots + \mathbf{v}_n$ son linealmente independientes.

49. Sea $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ un conjunto linealmente dependiente de vectores no nulos en un espacio vectorial V. Demuestre que al menos uno de los vectores de S puede ser expresado como combinación lineal de los vectores que lo preceden. Esto es, pruebe que existe un entero $k \leq n$ y escalares $a_1, a_2, \ldots, a_{k+1}$ tales que $\mathbf{v}_k = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_{k-1}\mathbf{v}_{k-1}$.

50. Sea $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ un conjunto de vectores con la propiedad de que el conjunto $\{\mathbf{v}_i, \mathbf{v}_j\}$ es linealmente dependiente si $i \neq j$. Demuestre que cada vector del conjunto es múltiplo de un solo vector en el conjunto.

□ 51. Sean f y g en $C'[0, 1]$. El *wronskiano** de f y g se define por

$$W(f, g)(x) = \begin{vmatrix} f(x) & g(x) \\ f'(x) & g'(x) \end{vmatrix}$$

Demuestre que si f y g son linealmente dependientes, entonces $W(f, g)(x) = 0$ para toda $x \in [0, 1]$.

□ 52. Determine una adecuada definición del wronskiano de las funciones $f_1, f_2, \ldots, f_n \in C^{(n-1)}[0, 1]$†

53. Supóngase que **u**, **v**, **w** son linealmente independientes. Demuestre o refute: **u** + **v**, **u** + **w**, y **v** + **w** son linealmente independientes.

54. ¿Para qué valores reales c, los vectores $(1 - c, 1 + c)$, $(1 + c, 1 - c)$ son linealmente independientes?

55. Demuestre que los vectores $(1, a, a^2)$, $(1, b, b^2)$ y $(1, c, c^2)$ son linealmente independientes si $a \neq b$, $a \neq c$ y $b \neq c$.

4.5 Combinación lineal y generación de espacio

Definición 1 **Combinación lineal.** Sean $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ vectores en un espacio vectorial V. Entonces, toda expresión de la forma

$$a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_n\mathbf{v}_n \qquad (1)$$

en donde $a_1, a_2, \ldots, a_n$ son escalares, se llama *combinación lineal* de $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$.

Ejemplo 1 En $\mathbb{R}^3$, $\begin{pmatrix}-7\\7\\7\end{pmatrix}$ es una combinación lineal de $\begin{pmatrix}-1\\2\\4\end{pmatrix}$ y $\begin{pmatrix}5\\-3\\1\end{pmatrix}$ ya que $\begin{pmatrix}-7\\7\\7\end{pmatrix} = 2\begin{pmatrix}-1\\2\\4\end{pmatrix} - \begin{pmatrix}5\\-3\\1\end{pmatrix}$.

Ejemplo 2 En M_{23}, $\begin{pmatrix}-3 & 2 & 8\\-1 & 9 & 3\end{pmatrix} = 3\begin{pmatrix}-1 & 0 & 4\\1 & 1 & 5\end{pmatrix} + 2\begin{pmatrix}0 & 1 & -2\\-2 & 3 & -6\end{pmatrix}$, lo cual demuestra que $\begin{pmatrix}-3 & 2 & 8\\-1 & 9 & 3\end{pmatrix}$ es una combinación lineal de $\begin{pmatrix}-1 & 0 & 4\\1 & 1 & 5\end{pmatrix}$ y de $\begin{pmatrix}0 & 1 & -2\\-2 & 3 & -6\end{pmatrix}$.

* Recibe este nombre en honor del matemático polaco Józef Maria Hoene-Wroński (1778-1853). Hoene-Wroński vivió la mayor parte de su vida adulta en Francia. Trabajó en la teoría de determinantes y fue conocido también por sus escritos críticos en la filosofía de las matemáticas.

† $C^{(n-1)}[0, 1]$ es el conjunto de funciones cuyas $(n-1)$-ésimas derivadas están definidas y son continuas en $[0, 1]$.

Ejemplo 3 En P_n, todo polinomio puede ser expresado como combinación lineal de los "monomios" $1, x, x^2, \ldots, x^n$.

Definición 2 **Generación de un espacio vectorial.** Los vectores $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ en un espacio vectorial V se dice que *generan* V, si todo vector en V puede expresarse como combinación lineal de ellos. Esto es, para todo $\mathbf{v} \in V$, existen escalares $a_1, a_2, \ldots, a_n$ tales que

$$\mathbf{v} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_n\mathbf{v}_n \qquad (2)$$

Ejemplo 4 Vimos en la Sección 3.1 que los vectores $\mathbf{i} = \begin{pmatrix}1\\0\end{pmatrix}$ y $\mathbf{j} = \begin{pmatrix}0\\1\end{pmatrix}$ generan $\mathbb{R}^2$. En la Sección 3.3 se vio que $\mathbf{i} = \begin{pmatrix}1\\0\\0\end{pmatrix}$, $\mathbf{j} = \begin{pmatrix}0\\1\\0\end{pmatrix}$ y $\mathbf{k} = \begin{pmatrix}0\\0\\1\end{pmatrix}$ generan $\mathbb{R}^3$.

De hecho, no es difícil obtener conjuntos de vectores que generen $\mathbb{R}^n$, como lo ilustra el siguiente teorema.

Teorema 1 Un conjunto de n vectores linealmente independientes en $\mathbb{R}^n$ genera a $\mathbb{R}^n$.

Demostración Sea $\mathbf{v}_1 = \begin{pmatrix}a_{11}\\a_{21}\\\vdots\\a_{n1}\end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix}a_{12}\\a_{22}\\\vdots\\a_{n2}\end{pmatrix}$, $\ldots$, $\mathbf{v}_n = \begin{pmatrix}a_{1n}\\a_{2n}\\\vdots\\a_{nn}\end{pmatrix}$ y

sea $\mathbf{v} = \begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}$ un vector en $\mathbb{R}^n$. Debemos demostrar que existen escalares $c_1, c_2, \ldots, c_n$ tales que

$$\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n$$

Esto es,

$$\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix} = c_1\begin{pmatrix}a_{11}\\a_{21}\\\vdots\\a_{n1}\end{pmatrix} + c_2\begin{pmatrix}a_{12}\\a_{22}\\\vdots\\a_{n2}\end{pmatrix} + \cdots + c_n\begin{pmatrix}a_{1n}\\a_{2n}\\\vdots\\a_{nn}\end{pmatrix} \qquad (3)$$

En (3), efectuamos las multiplicaciones, sumamos e igualamos componentes para obtener un sistema de n ecuaciones en las n incógnitas $c_1, c_2, \ldots, c_n$:

$$\begin{aligned} a_{11}c_1 + a_{12}c_2 + \cdots + a_{1n}c_n &= x_1 \\ a_{21}c_1 + a_{22}c_2 + \cdots + a_{2n}c_n &= x_2 \\ \vdots \\ a_{n1}c_1 + a_{n2}c_2 + \cdots + a_{nn}c_n &= x_n \end{aligned} \tag{4}$$

Se expresa (4) como $A\mathbf{c} = \mathbf{v}$, en donde

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \quad \text{y} \quad \mathbf{c} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$$

Pero el sistema (4) posee una solución única si y sólo si $\det A \neq 0$. Pero $\det A \neq 0$ porque las columnas de A son linealmente independientes. (Todo esto es consecuencia del Teorema 4.4.6.) Así pues, existe un solo vector $\mathbf{c}$ que satisface el sistema (4) y el teorema queda demostrado. ■

Observación. Este teorema no sólo muestra que $\mathbf{v}$ puede ser expresado como una combinación lineal de los vectores linealmente independientes $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$, sino que esto sólo es posible de *una sola manera* (ya que el vector solución $\mathbf{c}$ es único).

Ejemplo 5 Los vectores (2, −1, 4), (1, 0, 2) y (3, −1, 5) generan $\mathbb{R}^3$ pues $\begin{vmatrix} 2 & 1 & 3 \\ -1 & 0 & -1 \\ 4 & 2 & 5 \end{vmatrix} = -1 \neq 0$, así que son linealmente independientes.

Examinaremos a continuación brevemente conjuntos generadores de algunos otros espacios.

Ejemplo 6 Del Ejemplo 3, se deduce que los monomios $1, x, x^2, \ldots, x^n$ generan P_n.

Ejemplo 7 Como $\begin{pmatrix} a & b \\ c & d \end{pmatrix} = a\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + b\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + c\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + d\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ vemos que $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, y $\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ generan M_{22}.

Ejemplo 8 Sea P el espacio vectorial de los polinomios de cualquier grado. Entonces no existe conjunto *finito* alguno que genere a P. Para ver esto, supongamos que $p_1, p_2, \ldots, p_m$ son polinomios. Sea p_k el polinomio de más alto orden en este conjunto y sea $N = \deg p_k$. Entonces es claro que el polinomio $p(x) = x^{N+1}$ no puede ser escrito como combinación lineal de $p_1, p_2, \ldots, p_m$.

Ahora se considerará otro modo de encontrar subespacios de un espacio vectorial V.

Definición 3 **Espacio generado por un conjunto de vectores.** Sean $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ n vectores en un espacio vectorial V. El *espacio generado* por $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ es el conjunto de las combinaciones lineales de $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$. Esto es,

$$\text{gen}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\} = \{\mathbf{v}: \quad \mathbf{v} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_n\mathbf{v}_n\} \tag{5}$$

donde $a_1, a_2, \ldots, a_n$ son escalares.

Teorema 2 El gen $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ es un subespacio de V.

Demostración Es fácil y se deja como ejercicio (vea Problema 16). ■

Ejemplo 9 Sean $\mathbf{v}_1 = (2, -1, 4)$, $\mathbf{v}_2 = (4, 1, 6)$. Entonces $H = \text{gen}\{\mathbf{v}_1, \mathbf{v}_2\} = \{\mathbf{v}: \mathbf{v} = a_1(2, -1, 4) + a_2(4, 1, 6)\}$. ¿A qué se parece H? Si $\mathbf{v} = (x, y, z) \in H$, entonces $x = 2a_1 + 4a_2$, $y = -a_1 + a_2$, y $z = 4a_1 + 6a_2$. Si consideramos que (x, y, z) es fijo, entonces estas ecuaciones pueden verse como un sistema de tres ecuaciones en dos incógnitas a_1, a_2. Se resuelve el sistema en la forma usual:

$$\left(\begin{array}{cc|c} -1 & 1 & y \\ 2 & 4 & x \\ 4 & 6 & z \end{array}\right) \xrightarrow{M_1(-1)} \left(\begin{array}{cc|c} 1 & -1 & -y \\ 2 & 4 & x \\ 4 & 6 & z \end{array}\right) \xrightarrow[A_{1,3}(-4)]{A_{1,2}(-2)} \left(\begin{array}{cc|c} 1 & -1 & -y \\ 0 & 6 & x+2y \\ 0 & 10 & z+4y \end{array}\right)$$

$$\xrightarrow{M_2(\frac{1}{6})} \left(\begin{array}{cc|c} 1 & -1 & -y \\ 0 & 1 & (x+2y)/6 \\ 0 & 10 & z+4y \end{array}\right) \xrightarrow[A_{2,3}(-10)]{A_{2,1}(1)} \left(\begin{array}{cc|c} 1 & 0 & x/6-2y/3 \\ 0 & 1 & x/6+y/3 \\ 0 & 0 & -5x/3+2y/3+z \end{array}\right)$$

Del Capítulo 1, vemos que el sistema posee solución si $-5x/3 + 2y/3 + z = 0$; o, multiplicando por -3, si

$$5x - 2y - 3z = 0 \tag{6}$$

Pero la Ecuación (6) es la ecuación de un plano en $\mathbb{R}^3$ que pasa por el origen.

El último ejemplo puede ser generalizado para demostrar el hecho interesante que sigue:

> *El espacio generado por dos vectores no nulos en $\mathbb{R}^3$ y que no sean paralelos es un plano que pasa por el origen.*

Para una demostración sugerida, véanse los Problemas 19 y 20.

Es posible dar una interpretación geométrica de este resultado. Consideremos los vectores de la Figura 4.1. Conocemos (de la Sección 3.1) la interpretación geométrica de los vectores $2\mathbf{u}$, $-\mathbf{u}$ y $\mathbf{u} + \mathbf{v}$, por vía de ejemplo. Usando éstos se ve que cualquier otro vector en el plano de $\mathbf{u}$ y de $\mathbf{v}$ se puede obtener como combinación lineal de $\mathbf{u}$ y de $\mathbf{v}$. La Figura 4.2 muestra cómo en cuatro situaciones diferentes, un vector $\mathbf{w}$ ubicado en el plano de $\mathbf{u}$ y $\mathbf{v}$ puede ser escrito como $\alpha\mathbf{u} + \beta\mathbf{v}$ eligiendo apropiadamente los números α y β.

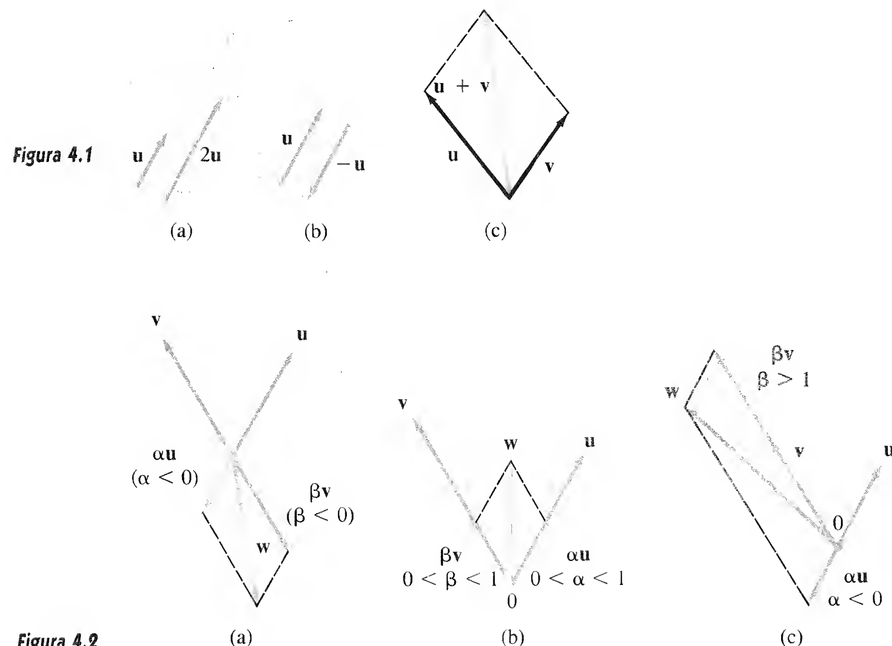


Figura 4.1

Figura 4.2

Concluiremos esta sección citando un resultado útil. Su demostración no es difícil y se deja como ejercicio (vea Problema 21).

Teorema 3 Sean los $n + 1$ vectores $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n, \mathbf{v}_{n+1}$, de un espacio vectorial V. Si $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ generan V, entonces $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n, \mathbf{v}_{n+1}$ también generan V. Esto es, la adición de uno (o más) vectores a un conjunto generador resulta en otro conjunto generador.

Problemas 4.5

En los Problemas del 1 *al* 13 *determinar si el conjunto de vectores dado genera al espacio vectorial que se da.*

1. En $\mathbb{R}^2$: $\begin{pmatrix}1\\2\end{pmatrix}, \begin{pmatrix}3\\4\end{pmatrix}$

2. En $\mathbb{R}^2$: $\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}2\\1\end{pmatrix}, \begin{pmatrix}2\\2\end{pmatrix}$

3. En $\mathbb{R}^2$: $\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}2\\2\end{pmatrix}, \begin{pmatrix}5\\5\end{pmatrix}$

4. En $\mathbb{R}^3$: $\begin{pmatrix}1\\2\\3\end{pmatrix}, \begin{pmatrix}-1\\2\\3\end{pmatrix}, \begin{pmatrix}5\\2\\3\end{pmatrix}$

5. En $\mathbb{R}^3$: $\begin{pmatrix}1\\1\\1\end{pmatrix}, \begin{pmatrix}0\\1\\1\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix}$

6. En $\mathbb{R}^3$: $\begin{pmatrix}2\\0\\1\end{pmatrix}, \begin{pmatrix}3\\1\\2\end{pmatrix}, \begin{pmatrix}1\\1\\1\end{pmatrix}, \begin{pmatrix}7\\3\\5\end{pmatrix}$

7. En $\mathbb{R}^3$: $(1, -1, 2), (1, 1, 2), (0, 0, 1)$

8. En $\mathbb{R}^3$: $(1, -1, 2), (-1, 1, 2), (0, 0, 1)$

9. En P_2: $1-x, 3-x^2$

10. En P_2: $1-x, 3-x^2, x$

11. En M_{22}: $\begin{pmatrix}2 & 1\\0 & 0\end{pmatrix}, \begin{pmatrix}0 & 0\\2 & 1\end{pmatrix}, \begin{pmatrix}3 & -1\\0 & 0\end{pmatrix}, \begin{pmatrix}0 & 0\\3 & 1\end{pmatrix}$

12. En M_{22}: $\begin{pmatrix}1 & 0\\1 & 0\end{pmatrix}, \begin{pmatrix}1 & 2\\0 & 0\end{pmatrix}, \begin{pmatrix}4 & -1\\3 & 0\end{pmatrix}, \begin{pmatrix}-2 & 5\\6 & 0\end{pmatrix}$.

13. En M_{23}: $\begin{pmatrix}1 & 0 & 0\\0 & 0 & 0\end{pmatrix}, \begin{pmatrix}0 & 1 & 0\\0 & 0 & 0\end{pmatrix}, \begin{pmatrix}0 & 0 & 1\\0 & 0 & 0\end{pmatrix}$, $\begin{pmatrix}0 & 0 & 0\\1 & 0 & 0\end{pmatrix}, \begin{pmatrix}0 & 0 & 0\\0 & 1 & 0\end{pmatrix}, \begin{pmatrix}0 & 0 & 0\\0 & 0 & 1\end{pmatrix}$

14. Demuestre que P_2 no puede ser generado por dos polinomios.

★ **15.** Si $p_1, p_2, \ldots, p_m$ generan P_n, entonces $m \geq n + 1$.

16. Demuestre que si $\mathbf{u}$, $\mathbf{v}$ están en gen $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$, entonces $\mathbf{u} + \mathbf{v}$ y $\alpha\mathbf{u}$ también lo están. [*Sugerencia:* Utilizar la definición de espacio generado para expresar $\mathbf{u} + \mathbf{v}$ y $\alpha\mathbf{u}$ como combinaciones lineales de $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$.]

17. Demuestre que el conjunto infinito $\{1, x, x^2, x^3, \ldots\}$ genera P, el espacio vectorial de polinomios.

18. Considere que H es un espacio de V que contiene a $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$. Demuestre que gen $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\} \subseteq H$. Es decir, gen $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ es el *más pequeño* subespacio de V que contiene a $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$.

19. Sean $\mathbf{v}_1 = (x_1, y_1, z_1)$ y $\mathbf{v}_2 = (x_2, y_2, z_2)$ vectores en $\mathbb{R}^3$. Demuestre que si $\mathbf{v}_2 = c\mathbf{v}_1$, entonces gen $\{\mathbf{v}_1, \mathbf{v}_2\}$ es una recta que pasa por el origen.

★ ★ **20.** En el Problema 19, supóngase que $\mathbf{v}_1$ y $\mathbf{v}_2$ son independientes. Demuéstrese que $H = \text{gen}\{\mathbf{v}_1, \mathbf{v}_2\}$ es un plano que pasa por el origen. ¿Cuál es la ecuación de este plano? [*Sugerencia:* Si $(x, y, z) \in H$, escriba $\mathbf{v} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2$ y encuentre una condición que relacione x, y y z de modo que el sistema 3×2 resultante tenga una solución.]

21. Demuestre el Teorema 3. [*Sugerencia:* Si $\mathbf{v} \in V$, escriba $\mathbf{v}$ como combinación lineal de $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n, \mathbf{v}_{n+1}$, con el coeficiente de $\mathbf{v}_{n+1}$ igual a cero.]

22. Demuestre que M_{22} puede ser generado por matrices invertibles.

★ **23.** Sean $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ y $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ $2n$ vectores en un espacio vectorial V. Suponga que

$$\begin{aligned}\mathbf{v}_1 &= a_{11}\mathbf{u}_1 + a_{12}\mathbf{u}_2 + \cdots + a_{1n}\mathbf{u}_n\\ \mathbf{v}_2 &= a_{21}\mathbf{u}_1 + a_{22}\mathbf{u}_2 + \cdots + a_{2n}\mathbf{u}_n\\ &\vdots\\ \mathbf{v}_n &= a_{n1}\mathbf{u}_1 + a_{n2}\mathbf{u}_2 + \cdots + a_{nn}\mathbf{u}_n\end{aligned}$$

Demuestre que si

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & & \vdots\\ a_{n1} & a_{n2} & \cdots & a_{nn}\end{vmatrix} \neq 0$$

entonces $\text{gen}\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\} = \text{gen}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$.

24. Sea $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ un conjunto linealmente independiente y supóngase que $\mathbf{v} \notin \text{gen}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$. Demuestre que $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n, \mathbf{v}\}$ es un conjunto linealmente independiente.

25. Halle un conjunto de tres vectores linealmente independientes en $\mathbb{R}^3$ que contenga los vectores $\begin{pmatrix}2\\1\\2\end{pmatrix}$ y $\begin{pmatrix}-1\\3\\4\end{pmatrix}$. $\left[\textit{Sugerencia:}\text{ Realice la obtención de un vector } \mathbf{v} \notin \text{gen}\left\{\begin{pmatrix}2\\1\\2\end{pmatrix}, \begin{pmatrix}-1\\3\\4\end{pmatrix}\right\}.\right]$

26. Determine un conjunto de tres vectores linealmente independientes en P_2 tal que contenga a los polinomios $1 - x^2$ y $1 + x^2$.

4.6 Base y dimensión

Hemos visto que en $\mathbb{R}^2$ es conveniente escribir los vectores en términos de los vectores $\mathbf{i} = \begin{pmatrix}1\\0\end{pmatrix}$ y $\mathbf{j} = \begin{pmatrix}0\\1\end{pmatrix}$. En $\mathbb{R}^3$ se expresan los vectores en función de $\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix}$. Ahora generalizaremos esta idea.

Definición 1 **Base.** Un conjunto de vectores $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ forma una *base* para V si

i. $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ es linealmente independiente.
ii. $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ genera V.

Hemos visto ya algunos ejemplos de bases. En el Teorema 4.5.1, por ejemplo, se vio que cualquier conjunto de n vectores linealmente independientes en $\mathbb{R}^n$, genera $\mathbb{R}^n$. Así pues,

> *Todo conjunto de n vectores linealmente independientes en $\mathbb{R}^n$ es una base en $\mathbb{R}^n$.*

En $\mathbb{R}^n$ definimos

$$\mathbf{e}_1 = \begin{pmatrix}1\\0\\0\\\vdots\\0\end{pmatrix}, \mathbf{e}_2 = \begin{pmatrix}0\\1\\0\\\vdots\\0\end{pmatrix}, \mathbf{e}_3 = \begin{pmatrix}0\\0\\1\\\vdots\\0\end{pmatrix}, \ldots, \mathbf{e}_n = \begin{pmatrix}0\\0\\0\\\vdots\\1\end{pmatrix}$$

Como los términos $\mathbf{e}_i$ son las columnas de la matriz identidad (cuyo determinante es 1), entonces $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ es linealmente independiente y, por tanto, constituye una base en $\mathbb{R}^n$. Esta entidad especial se llama *base canónica* en $\mathbb{R}^n$. Encontraremos ahora bases para otros espacios.

Ejemplo 1 Por el Ejemplo 4.4.8, los polinomios $1, x, x^2, x^3$ son linealmente independientes en P_3. Por el Ejemplo 4.5.3, estos polinomios generan P_3. Así pues, $\{1, x, x^2, x^3\}$ es una base de P_3. En general, los monomios $\{1, x, x^2, x^3, \ldots, x^n\}$ constituyen una base de P_n. A ésta se le llama *base canónica* de P_n.

Ejemplo 2 Se vio en el Ejemplo 4.5.7, que $\begin{pmatrix}1&0\\0&0\end{pmatrix}, \begin{pmatrix}0&1\\0&0\end{pmatrix}, \begin{pmatrix}0&0\\1&0\end{pmatrix}$ y $\begin{pmatrix}0&0\\0&1\end{pmatrix}$ generan M_{22}. Si $c_1\begin{pmatrix}1&0\\0&0\end{pmatrix} + c_2\begin{pmatrix}0&1\\0&0\end{pmatrix} + c_3\begin{pmatrix}0&0\\1&0\end{pmatrix} + c_4\begin{pmatrix}0&0\\0&1\end{pmatrix} = \begin{pmatrix}0&0\\0&0\end{pmatrix}$, entonces, obviamente, $c_1 = c_2 = c_3 = c_4 = 0$. Así pues, éstas cuatro matrices son linealmente independientes y forman una base de M_{22}, que recibe el nombre de *base canónica* de M_{22}.

Ejemplo 3 Encontrar una base para el conjunto de vectores en el plano

$$\pi = \left\{\begin{pmatrix}x\\y\\z\end{pmatrix}: 2x - y + 3z = 0\right\}$$

Solución En el Ejemplo 4.2.6 vimos que π es un espacio vectorial. Para encontrar una base, nótese que si x y z se eligen arbitrariamente y $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \pi$, entonces $y = 2x + 3z$. Así que los vectores en π son de la forma $\begin{pmatrix} x \\ 2x+3z \\ z \end{pmatrix}$. Como x y z son arbitrarios, elegimos algunos valores simples para ellos. Tomando $x = 1$, $z = 0$, obtenemos $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}$; y tomando $x = 0$, $z = 1$, obtenemos $\mathbf{v}_2 = \begin{pmatrix} 0 \\ 3 \\ 1 \end{pmatrix}$. Entonces $\begin{pmatrix} x \\ 2x+3z \\ z \end{pmatrix} = x\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} + z\begin{pmatrix} 0 \\ 3 \\ 1 \end{pmatrix}$. De manera que $\{\mathbf{v}_1, \mathbf{v}_2\}$ generan π y, como obviamente son linealmente independientes (porque no son múltiplos el uno del otro) forman una base de π.

Si $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ forman una base de V, entonces cualquier otro vector $\mathbf{v} \in V$ puede ser escrito $\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n$. ¿Puede ser escrito lo anterior de otra manera como combinación lineal de las $\mathbf{v}_i$? La respuesta es *no*. (Vea la observación que sigue a la demostración del Teorema 4.5.1 para el caso $V = \mathbb{R}^n$.)

Teorema 1 Si $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ es una base de V y si $\mathbf{v} \in V$, entonces existe un conjunto *único* de escalares $c_1, c_2, \ldots, c_n$ tales que $\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_n\mathbf{v}_n$.

Demostración Al menos existe un conjunto tal de escalares, puesto que $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ genera a V. Supongamos que $\mathbf{v}$ puede ser expresado de dos modos como combinación lineal de los vectores de base. Esto es, supóngase que

$$\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n = d_1\mathbf{v}_1 + d_2\mathbf{v}_2 + \cdots + d_n\mathbf{v}_n$$

Entonces, restando, obtenemos la ecuación

$$(c_1 - d_1)\mathbf{v}_1 + (c_2 - d_2)\mathbf{v}_2 + \cdots + (c_n - d_n)\mathbf{v}_n = \mathbf{0}$$

Ahora, como los términos $\mathbf{v}_i$ son linealmente independientes, esta ecuación sólo puede satisfacerse si $c_1 - d_1 = c_2 - d_2 = \cdots = c_n - d_n = 0$. Así pues, $c_1 = d_1$, $c_2 = d_2$, $\ldots$, $c_n = d_n$ y el teorema queda demostrado. ■

Hemos visto que los espacios vectoriales pueden tener muchas bases. Surge naturalmente una pregunta: ¿Tienen todas las bases el mismo número de vectores? En $\mathbb{R}^3$, la respuesta es, ciertamente, sí. Para ver esto, nótese que cualesquiera tres vectores linealmente independientes de $\mathbb{R}^3$ forman una base. Menos de tres vectores no pueden formar una base, pues como vimos en la Sección 4.5, el espacio generado por dos vectores de $\mathbb{R}^3$ linealmente independientes, es un plano en $\mathbb{R}^3$ —y un plano no es todo $\mathbb{R}^3$. Similarmente, un conjunto de cuatro o más vectores en $\mathbb{R}^3$ no puede ser linealmente independiente, pues si los primeros tres vectores del conjunto fuesen linealmente independientes, formarían una base, y así todos los demás vectores podrían ser expresados como combinación lineal de ellos. De modo que todas las bases en $\mathbb{R}^3$ contienen tres vectores. El teorema siguiente dice que la pregunta formulada con anterioridad tiene por respuesta *sí* para todos los espacios vectoriales.

Teorema 2 Si $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\}$ y $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ son bases del espacio vectorial V, entonces $m = n$; esto es, cualesquiera dos bases en un espacio vectorial V poseen el mismo número de vectores.

Demostración* Sean $S_1 = \{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ y $S_2 = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ dos bases de V. Debemos demostrar que $m = n$. Se realizará lo anterior probando que si $m > n$, entonces S_1 es un conjunto linealmente dependiente, lo cual contradice la hipótesis de que S_1 es una base. Esto prueba que $m \le n$. La misma demostración hará ver que $n \le m$, y esto corrobora el teorema. Así pues, basta demostrar que si $m > n$ entonces S_1 es linealmente dependiente. Como S_2 es una base, podemos escribir cada $\mathbf{u}_i$ como combinación lineal de los términos $\mathbf{v}_i$. Se tiene que

$$\begin{aligned} \mathbf{u}_1 &= a_{11}\mathbf{v}_1 + a_{12}\mathbf{v}_2 + \cdots + a_{1n}\mathbf{v}_n \\ \mathbf{u}_2 &= a_{21}\mathbf{v}_1 + a_{22}\mathbf{v}_2 + \cdots + a_{2n}\mathbf{v}_n \\ &\vdots \\ \mathbf{u}_m &= a_{m1}\mathbf{v}_1 + a_{m2}\mathbf{v}_2 + \cdots + a_{mn}\mathbf{v}_n \end{aligned} \tag{1}$$

Para demostrar que S_1 es dependiente, hay que encontrar escalares $c_1, c_2, \ldots, c_m$, no todos cero, tales que

$$c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_m\mathbf{u}_m = \mathbf{0} \tag{2}$$

Sustituyendo (1) en (2),

$$\begin{aligned} &c_1(a_{11}\mathbf{v}_1 + a_{12}\mathbf{v}_2 + \cdots + a_{1n}\mathbf{v}_n) + c_2(a_{21}\mathbf{v}_1 + a_{22}\mathbf{v}_2 + \cdots + a_{2n}\mathbf{v}_n) \\ &\quad + \cdots + c_m(a_{m1}\mathbf{v}_1 + a_{m2}\mathbf{v}_2 + \cdots + a_{mn}\mathbf{v}_n) = \mathbf{0} \end{aligned} \tag{3}$$

que puede expresarse como

$$\begin{aligned} &(a_{11}c_1 + a_{21}c_2 + \cdots + a_{m1}c_m)\mathbf{v}_1 + (a_{12}c_1 + a_{22}c_2 + \cdots + a_{m2}c_m)\mathbf{v}_2 \\ &\quad + \cdots + (a_{1n}c_1 + a_{2n}c_2 + \cdots + a_{mn}c_m)\mathbf{v}_n = \mathbf{0} \end{aligned} \tag{4}$$

Puesto que $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ son linealmente independientes debemos tener que

$$\begin{aligned} a_{11}c_1 + a_{21}c_2 + \cdots + a_{m1}c_m &= 0 \\ a_{12}c_1 + a_{22}c_2 + \cdots + a_{m2}c_m &= 0 \\ &\vdots \\ a_{1n}c_1 + a_{2n}c_2 + \cdots + a_{mn}c_m &= 0 \end{aligned} \tag{5}$$

---

* Esta demostración se da para espacios vectoriales con bases que contienen un número finito de vectores. También tratamos los escalares como si fuesen números reales. Sin embargo, la demostración sirve también en el caso complejo.

El sistema (5) es uno homogéneo de n ecuaciones en m incógnitas $c_1, c_2, \ldots, c_m$. Y como $m > n$, el Teorema 1.7.1, expresa que el sistema posee un número infinito de soluciones. En consecuencia, existen escalares $c_1, c_2, \ldots, c_m$ no todos cero, tales que se satisface (2), y por tanto, S_1 es dependiente. Esta contradicción demuestra que $m \leq n$, y que al intercambiar los papeles de S_1 y de S_2, es posible demostrar que $n \leq m$ y la demostración queda completa. ■

Con este teorema estamos en condiciones de definir uno de los conceptos centrales del álgebra lineal.

Definición 2 **Dimensión.** Si el espacio vectorial V posee una base finita, la *dimensión* de V es el número de vectores en la base, y V se llama *espacio vectorial de dimensión finita*. De otra manera, V se denomina *espacio vectorial de dimensión infinita*. Si $V = \{\mathbf{0}\}$, entonces V se dice que es de *dimensión cero*.

Notación Se simboliza la dimensión de V como dim V.

Observación. No hemos demostrado que todo espacio vectorial posee una base. Esta difícil demostración aparece en la Sección 4.12. No necesitamos este hecho para que tenga sentido la Definición 2, pues *si* V tiene una base finita, entonces V es finito-dimensional. En caso contrario, V es de dimensión infinita. Así pues, para demostrar que V es de dimensión infinita, sólo basta probar que V no posee una base finita. Es posible lograr esto demostrando que V contiene un número infinito de vectores linealmente independientes (ver el Ejemplo 7 más adelante). No es necesario constituir una base infinita para V.

Ejemplo 4 Como n vectores linealmente independientes de $\mathbb{R}^n$ constituyen una base,

$$\dim \mathbb{R}^n = n$$

Ejemplo 5 Por el Ejemplo 1 y el Problema 4.4.47, los polinomios $\{1, x, x^2, \ldots, x^n\}$ constituyen una base de P_n. Así pues $\dim P_n = n + 1$.

Ejemplo 6 En M_{mn}, sea A_{ij} la matriz de $m \times n$ con el valor 1 en la posición ij, y con el valor cero en las demás posiciones. Es fácil demostrar que A_{ij} para $i = 1, 2, \ldots, m$ y $j = 1, 2, \ldots, n$, forman una base de M_{mn}. Así pues, $\dim M_{mn} = mn$.

Ejemplo 7 En el Ejemplo 4.5.8, se vio que P no es generado por un conjunto finito de polinomios. Por lo mismo, P no posee una base finita y es, por lo tanto, un espacio vectorial de dimensión infinita.

Existen algunos teoremas que expresan algo acerca de la dimensión de un espacio vectorial.

Teorema 3 Supóngase que dim $V = n$. Si $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m$ es un conjunto de m vectores linealmente independientes en V, entonces $m \leq n$.

Demostración Sea $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ una base de V. Si $m > n$, entonces, como en la demostración del Teorema 2, podemos encontrar constantes $c_1, c_2, \ldots, c_m$, no todas nulas, tales que se satisface la Ecuación (2). Esto contradiría la independencia lineal de los términos $\mathbf{u}_i$. Así pues, $m \leq n$. ■

Teorema 4 Sea H un subespacio del espacio vectorial V de dimensión finito. Entonces H es finito-dimensional y

$$\dim H \leq \dim V \qquad (6)$$

Demostración Sea dim $V = n$. Cualquier conjunto de vectores en H linealmente independiente, lo es también en V. Por el Teorema 3, cualquier conjunto linealmente independiente en H, cuando más, contiene n vectores. Así pues, H es de dimensión finita. Más aún, como una base en H es un conjunto linealmente independiente, se ve que $\dim H \leq n$. ■

El Teorema 4 tiene interesantes consecuencias. Dos de ellas se presentan a continuación.

□ **Ejemplo 8** Sea $P[0, 1]$ el conjunto de polinomios sobre el intervalo $[0, 1]$. Entonces $P[0, 1] \subset C[0, 1]$. Si $C[0, 1]$ fuese finito-dimensional, también lo sería $P[0, 1]$. Por el Ejemplo 7, esto no es así. Por tanto, $C[0, 1]$ es de dimensión infinita. Similarmente, como $P[0, 1] \subset C'[0, 1]$ (ya que todo polinomio es diferenciable) entonces también $C'[0, 1]$ es infinito-dimensional. En general:

Todo espacio vectorial que contiene un subespacio de dimensión infinita es de dimensión infinita.

Ejemplo 9 Podemos utilizar el Teorema 4 para hallar *todos* los subespacios de $\mathbb{R}^3$. Sea H un subespacio de $\mathbb{R}^3$. Existen cuatro posibilidades: $H = \{\mathbf{0}\}$; $\dim H = 1$, $\dim H = 2$ y $\dim H = 3$. Si $\dim H = 3$, entonces H contiene una base de tres vectores $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ linealmente independientes en $\mathbb{R}^3$. Pero entonces $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ forman una base de $\mathbb{R}^3$. Así que $H = \text{gen}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\} = \mathbb{R}^3$. De aquí que la única forma de conseguir un subespacio *propio* de $\mathbb{R}^3$ es considerar que $\dim H = 1$ o que $\dim H = 2$. Si $\dim H = 1$, entonces H posee una base con un solo vector $\mathbf{v} = (a, b, c)$. Sea $\mathbf{x} \in H$. Entonces $\mathbf{x} = t(a, b, c)$ para algún número real t (pues (a, b, c) genera H). Si $\mathbf{x} = (x, y, z)$, entonces $x = at$, $y = bt$, $z = ct$. Esto representa la ecuación de una recta en $\mathbb{R}^3$ que pasa por el origen y tiene un vector direccional (a, b, c).

Supongamos ahora que $\dim H = 2$ y sea $\mathbf{v}_1 = (a_1, b_1, c_1)$ y $\mathbf{v}_2 = (a_2, b_2, c_2)$ una base de H. Si $\mathbf{x} = (x, y, z) \in H$, entonces existen números reales s, t tales que $\mathbf{x} = s\mathbf{v}_1 + t\mathbf{v}_2$, o bien $(x, y, z) = s(a_1, b_1, c_1) + t(a_2, b_2, c_2)$. Por tanto

$$\begin{aligned} x &= sa_1 + ta_2 \\ y &= sb_1 + tb_2 \\ z &= sc_1 + tc_2 \end{aligned} \tag{7}$$

Sea $\mathbf{v}_3 = (\alpha, \beta, \gamma) = \mathbf{v}_1 \times \mathbf{v}_2$. Por el Teorema 3.4.2, parte (vi), tenemos que $\mathbf{v}_3 \cdot \mathbf{v}_1 = 0$ y que $\mathbf{v}_3 \cdot \mathbf{v}_2 = 0$. Ahora determínese

$$\begin{aligned} \alpha x + \beta y + \gamma z &= \alpha(sa_1 + ta_2) + \beta(sb_1 + tb_2) + \gamma(sc_1 + tc_2) \\ &= (\alpha a_1 + \beta b_1 + \gamma c_1)s + (\alpha a_2 + \beta b_2 + \gamma c_2)t \\ &= (\mathbf{v}_3 \cdot \mathbf{v}_1)s + (\mathbf{v}_3 \cdot \mathbf{v}_2)t = 0 \end{aligned}$$

Así que si $(x, y, z) \in H$, entonces $\alpha x + \beta y + \gamma z = 0$, lo cual demuestra que H es un plano que pasa por el origen y tiene un vector normal $\mathbf{v}_3 = \mathbf{v}_1 \times \mathbf{v}_2$. Hemos demostrado que

> *Los únicos subespacios propios de $\mathbb{R}^3$ son conjuntos de vectores sobre rectas y sobre planos que pasan por el origen.*

Ejemplo 10 Sea A una matriz de $m \times n$ y sea $\mathbf{S} = \{\mathbf{x} \in \mathbb{R}^n; A\mathbf{x} = \mathbf{0}\}$. Sean $\mathbf{x}_1 \in \mathbf{S}$ y $\mathbf{x}_2 \in \mathbf{S}$; entonces $A(\mathbf{x}_1 + \mathbf{x}_2) = A\mathbf{x}_1 + A\mathbf{x}_2 = \mathbf{0} + \mathbf{0} = \mathbf{0}$ y $A(\alpha\mathbf{x}_1) = \alpha(A\mathbf{x}_1) = \alpha\mathbf{0} = \mathbf{0}$, así que S es un subespacio de $\mathbb{R}^n$ y $\dim S \le n$; S se denomina *espacio solución* del sistema homogéneo $A\mathbf{x} = \mathbf{0}$. También recibe el nombre de *núcleo* (o *kernel* o espacio nulo).

Ejemplo 11 Encontrar una base (y su dimensión) del espacio solución S del sistema homogéneo

$$\begin{aligned} x + 2y - z &= 0 \\ 2x - y + 3z &= 0 \end{aligned}$$

Solución Aquí $A = \begin{pmatrix} 1 & 2 & -1 \\ 2 & -1 & 3 \end{pmatrix}$. Como A es una matriz de 2×3, S es un subespacio de $\mathbb{R}^3$. Reduciendo, encontramos, en forma sucesiva,

$$\left(\begin{array}{ccc|c} 1 & 2 & -1 & 0 \\ 2 & -1 & 3 & 0 \end{array}\right) \xrightarrow{A_{1,2}(-2)} \left(\begin{array}{ccc|c} 1 & 2 & -1 & 0 \\ 0 & -5 & 5 & 0 \end{array}\right)$$

$$\xrightarrow{M_2(-\frac{1}{5})} \left(\begin{array}{ccc|c} 1 & 2 & -1 & 0 \\ 0 & 1 & -1 & 0 \end{array}\right) \xrightarrow{A_{2,1}(-2)} \left(\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 \end{array}\right)$$

Entonces $y = z$ y $x = -z$, así que todas las soluciones son de la forma $\begin{pmatrix} -z \\ z \\ z \end{pmatrix}$.

De modo que $\begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$ es una base de $\mathbf{S}$ y $\dim \mathbf{S} = 1$. Nótese que $\mathbf{S}$ es el conjunto de vectores sobre la recta $x = -t$, $y = t$, $z = t$.

Ejemplo 12 Encontrar una base para el espacio solución $\mathbf{S}$ del sistema

$$\begin{aligned} 2x - y + 3z &= 0 \\ 4x - 2y + 6z &= 0 \\ -6x + 3y - 9z &= 0 \end{aligned}$$

Solución Reduciendo por filas o renglones igual que antes, resulta

$$\left(\begin{array}{ccc|c} 2 & -1 & 3 & 0 \\ 4 & -2 & 6 & 0 \\ -6 & 3 & -9 & 0 \end{array}\right) \xrightarrow[A_{1,3}(3)]{A_{1,2}(-2)} \left(\begin{array}{ccc|c} 2 & -1 & 3 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

lo cual da la ecuación única $2x - y + 3z = 0$. $\mathbf{S}$ es un plano, y por el Ejemplo 3, una base está dada por $\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}$ y $\begin{pmatrix} 0 \\ 3 \\ 1 \end{pmatrix}$, así que $\dim \mathbf{S} = 2$. Nótese que se ha demostrado que toda solución al sistema homogéneo puede escribirse como

$$c_1\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} + c_2\begin{pmatrix} 0 \\ 3 \\ 1 \end{pmatrix}$$

Por ejemplo, si $c_1 = 2$ y $c_2 = -3$, resulta la solución

$$\mathbf{x} = 2\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} - 3\begin{pmatrix} 0 \\ 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ -9 \\ -3 \end{pmatrix} = \begin{pmatrix} 2 \\ -5 \\ -3 \end{pmatrix}$$

Antes de terminar con esta sección, demostraremos un resultado muy útil para encontrar bases en un espacio vectorial arbitrario. Hemos visto que n vectores de $\mathbb{R}^n$ linealmente independientes forman una base de $\mathbb{R}^n$. Esto es cierto para *cualquier* espacio vectorial de dimensión finita.

Teorema 5 Cualesquiera n vectores linealmente independientes en un espacio vectorial V de dimensión n, constituyen una base.

Demostración Sean $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ los n vectores. Si generan V son una base. Si no lo hacen, entonces existe un vector $\mathbf{u} \in V$ tal que $\mathbf{u} \notin \text{gen}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$. Esto significa que los $n + 1$ vectores $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n, \mathbf{u}$ son linealmente independientes. Para ver esto nótese que si

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n + c_{n+1}\mathbf{u} = 0 \tag{8}$$

entonces $c_{n+1} = 0$, pues si no fuese así, $\mathbf{u}$ podría escribirse como combinación lineal de $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ al dividir entre c_{n+1} y poner todos los términos de la

suma, excepto $\mathbf{u}$, al lado derecho de la ecuación. Pero si $c_{n+1} = 0$, entonces (8) es

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n = 0$$

lo cual significa que $c_1 = c_2 = \cdots = c_n = 0$ puesto que los términos $\mathbf{v}_i$ son linealmente independientes. Ahora, sea $W = \text{gen}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n, \mathbf{u}\}$. Entonces como todos los vectores puestos entre las llaves están en V, se tiene que W es un subespacio de V. Como $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n, \mathbf{u}$ son linealmente independientes forman una base para W. Así pues, $\dim W = n + 1$. Por el Teorema 4, $\dim W \leq n$. Esta contradicción demuestra que *no* existe $\mathbf{u} \in V$ tal que $\mathbf{u} \notin \text{gen}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$. Así pues, $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ generan todo V y constituyen una base de V. ■

Problemas 4.6

En los Problemas del 1 al 10, determine si el conjunto de vectores dado es una base del espacio vectorial correspondiente.

1. En P_2: $1 - x^2, x$
2. En P_2: $-3x, 1 + x^2, x^2 - 5$
3. En P_2: $x^2 - 1, x^2 - 2, x^2 - 3$
4. En P_3: $1, 1 + x, 1 + x^2, 1 + x^3$
5. En P_3: $3, x^3 - 4x + 6, x^2$
6. En M_{22}: $\begin{pmatrix} 3 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 3 & 2 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} -5 & 1 \\ 0 & 6 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & -7 \end{pmatrix}$
7. En M_{22}: $\begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & b \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ c & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & d \end{pmatrix}$, en donde $abcd \neq 0$
8. En M_{22}: $\begin{pmatrix} -1 & 0 \\ 3 & 1 \end{pmatrix}, \begin{pmatrix} 2 & 1 \\ 1 & 4 \end{pmatrix}, \begin{pmatrix} -6 & 1 \\ 5 & 8 \end{pmatrix}, \begin{pmatrix} 7 & -2 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$
9. $H = \{(x, y) \in \mathbb{R}^2: x + y = 0\}$; $(1, -1)$
10. $H = \{(x, y) \in \mathbb{R}^2: x + y = 0\}$; $(1, -1), (-3, 3)$
11. Halle una base en $\mathbb{R}^3$ para el conjunto de vectores en el plano $2x - y - z = 0$.
12. Halle una base en $\mathbb{R}^3$ para el conjunto de vectores en el plano $3x - 2y + 6z = 0$.
13. Obtenga una base en $\mathbb{R}^3$ para el conjunto de vectores en la recta $x/2 = y/3 = z/4$.
14. Obtenga una base en $\mathbb{R}^3$ para el conjunto de vectores en la recta $x = 3t$, $y = -2t$, $z = t$.
15. Demuestre que los únicos subespacios propios de $\mathbb{R}^2$ son rectas que pasan por el origen.
16. En $\mathbb{R}^4$, sea $H = \{(x, y, z, w): ax + by + cz + dw = 0\}$, en donde $abcd \neq 0$.
    a. Demuestre que H es un subespacio de $\mathbb{R}^4$.
    b. Halle una base de H.
    c. ¿Cuál es la dimensión de H?

★ 17. En $\mathbb{R}^n$, un *hiperplano* es un subespacio de dimensión $n - 1$. Si H es un hiperplano en $\mathbb{R}^n$, demuestre que

$$H = \{(x_1, x_2, \ldots, x_n): a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0\}$$

en donde $a_1, a_2, \ldots, a_n$ son números reales fijos, no todos iguales a cero.

18. En $\mathbb{R}^5$, halle una base para el hiperplano

$$H = \{(x_1, x_2, x_3, x_4, x_5): 2x_1 - 3x_2 + x_3 + 4x_4 - x_5 = 0\}$$

En los Problemas del 19 al 23, obtenga una base para el espacio solución del sistema homogéneo indicado.

19. $\begin{aligned} x - y &= 0 \\ -2x + 2y &= 0 \end{aligned}$
20. $\begin{aligned} x - 2y &= 0 \\ 3x + y &= 0 \end{aligned}$
21. $\begin{aligned} x - y - z &= 0 \\ 2x - y + z &= 0 \end{aligned}$
22. $\begin{aligned} x - 3y + z &= 0 \\ -2x + 2y - 3z &= 0 \\ 4x - 8y + 5z &= 0 \end{aligned}$
23. $\begin{aligned} 2x - 6y + 4z &= 0 \\ -x + 3y - 2z &= 0 \\ -3x + 9y - 6z &= 0 \end{aligned}$

24. Determine una base para D_3, el espacio de las matrices diagonales de 3×3. ¿Cuál es $\dim D_3$?
25. ¿Cuál es $\dim D_n$, el espacio de las matrices diagonales de $n \times n$?
26. Sea S_{nn} el espacio vectorial de las matrices simétricas de $n \times n$. Demuestre que S_{nn} es un subespacio de M_{nn} y que $\dim S_{nn} = [n(n + 1)]/2$.
27. Supóngase que $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ son vectores linealmente independientes de un espacio vectorial V de dimensión n, y $m < n$. Demuestre que el conjunto $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m\}$ puede ser ampliado a una base de V. Esto es, que existen vectores $\mathbf{v}_{m+1}, \mathbf{v}_{m+2}, \ldots, \mathbf{v}_n$ tales que $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ forman una base. [*Sugerencia:* Ver la demostración del Teorema 5.]
28. Sea $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ una base de V. Considérese que $\mathbf{u}_1 = \mathbf{v}_1$, $\mathbf{u}_2 = \mathbf{v}_1 + \mathbf{v}_2$, $\mathbf{u}_3 = \mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}_3, \ldots, \mathbf{u}_n = \mathbf{v}_1 + \mathbf{v}_2 + \cdots + \mathbf{v}_n$. Demuestre que $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ también es una base de V.
29. Pruebe que si $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ genera a V, entonces $\dim V \leq n$. [*Sugerencia:* Utilice el resultado del Problema 4.4.49.]
30. Sean H y K subespacios de V tales que $H \subseteq K$ y $\dim H = \dim K < \infty$. Demuestre que $H = K$.
31. Sean H y K subespacios de V y defínase $H + K = \{\mathbf{h} + \mathbf{k}: \mathbf{h} \in H \text{ y } \mathbf{k} \in K\}$.
    a. Pruebe que $H + K$ es un subespacio de V.
    b. Si $H \cap K = \{\mathbf{0}\}$, entonces $\dim (H + K) = \dim H + \dim K$.

★ 32. Si H es un subespacio del espacio vectorial finito-dimensional V, demuestre que existe un subespacio único K de V tal que (*a*) $H \cap K = \{\mathbf{0}\}$, (*b*) $H + K = V$.

33. Demuestre que dos vectores $\mathbf{v}_1$ y $\mathbf{v}_2$ de $\mathbb{R}^2$, con uno de sus extremos en el origen, son colineales si y sólo si $\dim \text{gen}\{\mathbf{v}_1, \mathbf{v}_2\} = 1$.
34. Demuestre que tres vectores $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$ en $\mathbb{R}^3$, con uno de sus extremos en el origen, son coplanares si y sólo si $\dim \text{gen}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\} \leq 2$.
35. Pruebe que cualesquiera n vectores que generan un espacio n-dimensional V forman una base de V. [*Sugerencia:* Demuestre que si los n vectores no son linealmente independientes, entonces $\dim V < n$.]

★ 36. Demuestre que todo subespacio de un espacio vectorial finito-dimensional posee una base.

37. Encuentre dos bases para $\mathbb{R}^4$ que contengan a los vectores $(1, 0, 1, 0)$ y $(0, 1, 0, 1)$, y no tengan otros vectores en común.
38. ¿Para qué valores del número real a constituyen una base de $\mathbb{R}^3$ los vectores $(a, 1, 0)$, $(1, 0, a)$ y $(1 + a, 1, a)$?

4.7 Rango, nulidad, espacio de renglones y espacio de columnas de una matriz

En la Sección 4.4 se presentó la noción de independencia lineal. Demostramos que si A es una matriz de $n \times n$, entonces los renglones y las columnas de A forman conjuntos de vectores linealmente independientes si y sólo si $\det A \neq 0$. Sin embargo, si $\det A = 0$, o si A no es cuadrada, entonces estos resultados no dicen nada acerca del número de renglones o columnas linealmente independientes de A. En esta sección subsanaremos esta omisión. También se demostrará como obtener una base para el espacio generado por un conjunto de vectores usando reducción por filas o renglones.

Sea A una matriz de $m \times n$, y

$$N_A = \{\mathbf{x} \in \mathbb{R}^n: \ A\mathbf{x} = \mathbf{0}\} \qquad (1)$$

Como se vio en el Ejemplo 4.6.10, N_A es un subespacio de $\mathbb{R}^n$.

Definición 1 ***Núcleo (o kernel) y nulidad de una matriz.*** N_A se llama *núcleo* (o *kernel*) de A, y al número $\nu(A) = \dim N_A$ se le denomina *nulidad* de A. Si N_A contiene sólo al vector cero, entonces $\nu(A) = 0$. (N_A se representa también por $\ker A$, o bien, por $\text{núcl} A$.)

Ejemplo 1 Sea $A = \begin{pmatrix} 1 & 2 & -1 \\ 2 & -1 & 3 \end{pmatrix}$. Entonces, como se vio en el Ejemplo 4.6.11, N_A está generado por $\begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$ y $\nu(A) = 1$.

Ejemplo 2 Sea $A = \begin{pmatrix} 2 & -1 & 3 \\ 4 & -2 & 6 \\ -6 & 3 & -9 \end{pmatrix}$. De manera que por el Ejemplo 4.6.12,

$\left\{\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 3 \\ 1 \end{pmatrix}\right\}$ es una base de N_A y $\nu(A) = 2$.

Teorema 1 Sea A una matriz de $n \times n$. Entonces A es invertible si y sólo si $\nu(A) = 0$.

Demostración Por el Teorema Resumen [Teorema 4.4.6, partes (*i*) y (*ii*)], A es invertible si y sólo si el sistema homogéneo $A\mathbf{x} = \mathbf{0}$ posee solamente la solución trivial $\mathbf{x} = \mathbf{0}$. De la Ecuación (1), esto significa que A es invertible si y sólo si $N_A = \{\mathbf{0}\}$. Así pues, A es invertible si y sólo si $\nu(A) = \dim N_A = 0$. ■

Definición 2 ***Imagen (o espacio imagen) de una matriz.*** Sea A una matriz de $m \times n$. Entonces la *imagen* de A, denotada por $\text{Imag} A$, está dada por

$$\text{Imag } A = \{\mathbf{y} \in \mathbb{R}^m: \ A\mathbf{x} = \mathbf{y} \ \text{ para } \mathbf{x} \in \mathbb{R}^n\} \qquad (2)$$

Teorema 2 Sea A una matriz de $m \times n$, entonces $\text{Imag} A$ es un subespacio de $\mathbb{R}^m$.

Demostración Supongamos que $\mathbf{y}_1$ y $\mathbf{y}_2$ están en $\text{Imag} A$. Entonces existen vectores $\mathbf{x}_1$, $\mathbf{x}_2$ en $\mathbb{R}^n$ tales que $\mathbf{y}_1 = A\mathbf{x}_1$, $\mathbf{y}_2 = A\mathbf{x}_2$. Por lo tanto,

$$A(\alpha\mathbf{x}_1) = \alpha A\mathbf{x}_1 = \alpha\mathbf{y}_1 \quad \text{y} \quad A(\mathbf{x}_1 + \mathbf{x}_2) = A\mathbf{x}_1 + A\mathbf{x}_2 = \mathbf{y}_1 + \mathbf{y}_2$$

así que $\alpha\mathbf{y}_1$ y $\mathbf{y}_1 + \mathbf{y}_2$ están en $\text{Imag} A$. Del Teorema 4.3.1, se sigue que $\text{Imag } A$ es un subespacio de $\mathbb{R}^m$. ■

Definición 3 ***Rango de una matriz.*** Sea A una matriz de $m \times n$. El *rango* de A, denotado por $\rho(A)$, está dado por

$$\rho(A) = \dim \text{Imag} A$$

Antes de exponer ejemplos, daremos dos definiciones y un teorema que facilitarán la determinación del rango.*

Definición 4 ***Espacios de renglones y de columnas de una matriz.*** Si A es una matriz de $m \times n$, sean $\{\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_m\}$ y $\{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n\}$ los conjuntos de renglones y de columnas de A, respectivamente. Entonces se define que

$$R_A = \textit{espacio de renglones} \text{ de } A = \text{gen}\{\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_m\} \qquad (3)$$

y

$$C_A = \textit{espacio de columnas} \text{ de } A = \text{gen}\{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n\} \qquad (4)$$

Nótese que R_A es un subespacio de $\mathbb{R}^n$ y que C_A lo es de $\mathbb{R}^m$.

Hemos introducido una gran cantidad de notaciones en estas últimas partes. Detengámonos un momento a ilustrar estas ideas con un ejemplo.

* **N. de R.T.** En esta área matemática, sobre todo, conviene tener presente siempre que el término español rango equivale exclusivamente a la palabra inglesa *rank*. La traducción errónea del inglés *range* como "rango" causaría aquí una grave confusión. De modo que *range* = imagen (en este campo matemático) y *rank* = rango.

Ejemplo 3 Sea $A = \begin{pmatrix} 1 & 2 & -1 \\ 2 & -1 & 3 \end{pmatrix}$. Entonces A es una matriz de 2×3.

i. El *kernel* o *núcleo* de A es $N_A = \{\mathbf{x} \in \mathbb{R}^3\colon A\mathbf{x} = \mathbf{0}\}$. Como se vio en el Ejemplo 1, $N_A = \text{gen}\left\{\begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}\right\}$.

ii. La *nulidad* de A es $\nu(A) = \dim N_A = 1$.

iii. La *imagen* de A es $\text{Imag}\, A = \{\mathbf{y} \in \mathbb{R}^2\colon A\mathbf{x} = \mathbf{y} \text{ para } \mathbf{x} \in \mathbb{R}^3\}$. Sea $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ en $\mathbb{R}^2$. Entonces, si $\mathbf{y} \in \text{Imag}\, A$, existe una $\mathbf{x} \in \mathbb{R}^3$ tal que $A\mathbf{x} = \mathbf{y}$. Escribiendo $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$, resulta

$$\begin{pmatrix} 1 & 2 & -1 \\ 2 & -1 & 3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

o bien

$$\begin{aligned} x_1 + 2x_2 - x_3 &= y_1 \\ 2x_1 - x_2 + 3x_3 &= y_2. \end{aligned}$$

Reduciendo este sistema por filas, obtenemos

$$\left(\begin{array}{ccc|c} 1 & 2 & -1 & y_1 \\ 2 & -1 & 3 & y_2 \end{array}\right) \xrightarrow{A_{1,2}(-2)} \left(\begin{array}{ccc|c} 1 & 2 & -1 & y_1 \\ 0 & -5 & 5 & y_2 - 2y_1 \end{array}\right)$$

$$\xrightarrow{M_2(-\frac{1}{5})} \left(\begin{array}{ccc|c} 1 & 2 & -1 & y_1 \\ 0 & 1 & -1 & \frac{2y_1 - y_2}{5} \end{array}\right) \xrightarrow{A_{2,1}(-2)} \left(\begin{array}{ccc|c} 1 & 0 & 1 & \frac{y_1 + 2y_2}{5} \\ 0 & 1 & -1 & \frac{2y_1 - y_2}{5} \end{array}\right)$$

Así que si elegimos x_3 en forma arbitraria, resulta

$$x_1 = -x_3 + \frac{y_1 + 2y_2}{5} \quad \text{y} \quad x_2 = x_3 + \frac{2y_1 - y_2}{5}$$

Esto es, para cada $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \in \mathbb{R}^2$, existe un número infinito de vectores $\mathbf{x} \in \mathbb{R}^3$ tales que $A\mathbf{x} = \mathbf{y}$. Así pues, $\text{Imag}\, A = \mathbb{R}^2$. Nótese, por ejemplo, que si $\mathbf{y} = \begin{pmatrix} 2 \\ -3 \end{pmatrix}$, entonces, eligiendo $x_3 = 0$ (la elección más simple), tenemos

$$x_1 = \frac{2 + 2(-3)}{5} = -\tfrac{4}{5} \quad \text{y} \quad x_2 = \frac{2(2) - (-3)}{5} = \tfrac{7}{5}$$

y

$$A\mathbf{x} = \begin{pmatrix} 1 & 2 & -1 \\ 2 & -1 & 3 \end{pmatrix}\begin{pmatrix} -\frac{4}{5} \\ \frac{7}{5} \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{10}{5} \\ -\frac{15}{5} \end{pmatrix} = \begin{pmatrix} 2 \\ -3 \end{pmatrix} = \mathbf{y}$$

iv. El *rango* de A es $\rho(A) = \dim \text{Imag}\, A = \dim \mathbb{R}^2 = 2$.

v. El *espacio de renglones* de A es $R_A = \text{gen}\{(1, 2, -1), (2, -1, 3)\}$. Como estos dos vectores son linealmente independientes, vemos que R_A es un subespacio bidimensional de $\mathbb{R}^3$. Del Ejemplo 4.6.9 observamos que R_A es un plano que pasa por el origen.

vi. El *espacio de columnas* de A es

$$C_A = \text{gen}\left\{\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ 3 \end{pmatrix}\right\} = \mathbb{R}^2$$

ya que $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ y $\begin{pmatrix} 2 \\ -1 \end{pmatrix}$, siendo linealmente independientes, constituyen una base de $\mathbb{R}^2$.

En el Ejemplo 3, podemos observar que $\text{Imag}\, A = C_A = \mathbb{R}^2$ y que $\dim R_A = \dim C_A = \dim \text{Imag}\, A = \rho(A) = 2$. Esto no es una coincidencia.

Teorema 3 Si A es una matriz de $m \times n$, entonces

i. $C_A = \text{Imag}\, A$

ii. $\dim R_A = \dim C_A = \dim \text{Imag}\, A = \rho(A)$

La demostración de este teorema no es difícil, pero sí laboriosa. La diferiremos hasta el final de la sección.

Ejemplo 4 Obtener una base de $\text{Imag}\, A$ y determinar el rango de $A = \begin{pmatrix} 2 & -1 & 3 \\ 4 & -2 & 6 \\ -6 & 3 & -9 \end{pmatrix}$.

Solución Como $\mathbf{r}_2 = 2\mathbf{r}_1$ y $\mathbf{r}_3 = -3\mathbf{r}_1$, vemos que $\rho(A) = 1$. Así pues, cualquier columna de C_A es una base de $C_A = \text{Imag}\, A$. Por ejemplo, $\begin{pmatrix} 2 \\ 4 \\ -6 \end{pmatrix}$ es una base de $\text{Imag}\, A$.

El siguiente teorema simplificará los cálculos.

Teorema 4 Si A es equivalente por filas o renglones (o por columnas) a B, entonces $R_A = R_B$, $\rho(A) = \rho(B)$ y $\nu(A) = \nu(B)$.

Demostración Recordamos de la Definición 1.8.3, que A es equivalente a B por filas, si A puede ser "reducida" a B por medio de operaciones elementales sobre filas. La definición relativa a "equivalente por columnas" es similar. Supongamos que C es la matriz obtenida al efectuar una operación por filas sobre A. Demostraremos primero que $R_A = R_C$. Como B se obtiene de A por medio de varias operaciones de este tipo, nuestro primer resultado, aplicado varias veces, implicará que $R_A = R_B$.

Caso 1: Intercambio de dos renglones de A. Entonces $R_A = R_C$ porque los renglones de A y de C son los mismos (sólo se escriben en un orden diferente).

Caso 2: Multiplicación del renglón i de A por $c \neq 0$. Si los renglones de A son $\{\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_i, \ldots, \mathbf{r}_m\}$, entonces las filas de C son $\{\mathbf{r}_1, \mathbf{r}_2, \ldots, c\mathbf{r}_i, \ldots, \mathbf{r}_m\}$. Obviamente, $c\mathbf{r}_i = c(\mathbf{r}_i)$ y $\mathbf{r}_i = (1/c)\mathbf{r}_i$. Así pues, cada renglón de C es un múltiplo de un renglón de A, y viceversa. Esto significa que cada renglón de C está en el espacio generado por los renglones de A y viceversa. Tenemos que

$$R_A \subseteq R_C \quad \text{y} \quad R_C \subseteq R_A, \text{ así que } R_C = R_A$$

Caso 3: Multiplicación de la fila i de A por $c \neq 0$, y su suma a la fila j. Ahora los renglones de C son $\{\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_i, \ldots, c\mathbf{r}_i + \mathbf{r}_j, \ldots, \mathbf{r}_m\}$. Aquí

$$\mathbf{r}_j = \underbrace{(c\mathbf{r}_i + \mathbf{r}_j)}_{\substack{\text{renglón } j\text{-ésimo} \\ \text{de } C}} - c\underset{\substack{\uparrow \\ \text{renglón } i\text{-ésimo} \\ \text{de } C}}{\mathbf{r}_i}$$

de modo que cada renglón puede ser escrito como combinación lineal de los renglones de C, y viceversa. Tenemos, como antes,

$$R_A \subseteq R_C \quad \text{y} \quad R_C \subseteq R_A \text{ así que } R_C = R_A$$

Hemos demostrado que $R_A = R_B$. De donde $\rho(R_A) = \rho(R_B)$. Finalmente, el conjunto de soluciones de $A\mathbf{x} = 0$ no cambia ante operaciones elementales por filas. En consecuencia, $N_A = N_B$, así que $\nu(A) = \nu(B)$. ■

Ejemplo 5 Determinar el rango y el espacio de renglones de $A = \begin{pmatrix} 1 & -1 & 3 \\ 2 & 0 & 4 \\ -1 & -3 & 1 \end{pmatrix}$.

Solución Reducimos por filas para obtener una matriz más simple:

$$\begin{pmatrix} 1 & -1 & 3 \\ 2 & 0 & 4 \\ -1 & -3 & 1 \end{pmatrix} \xrightarrow[A_{1,3}(1)]{A_{1,2}(-2)} \begin{pmatrix} 1 & -1 & 3 \\ 0 & 2 & -2 \\ 0 & -4 & 4 \end{pmatrix}$$

$$\xrightarrow{M_2(\frac{1}{2})} \begin{pmatrix} 1 & -1 & 3 \\ 0 & 1 & -1 \\ 0 & -4 & 4 \end{pmatrix} \xrightarrow{A_{2,3}(4)} \begin{pmatrix} 1 & -1 & 3 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix} = B$$

Como B posee dos renglones independientes, $\rho(B) = \rho(A) = 2$ y

$$R_A = \text{gen } \{(1, -1, 3), (0, 1, -1)\}$$

El Teorema 4 es útil cuando queremos encontrar una base para el espacio generado por un conjunto de vectores.

Ejemplo 6 Encontrar una base para el espacio generado por los vectores

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} -2 \\ 0 \\ 4 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} 0 \\ 4 \\ -2 \end{pmatrix}, \quad \text{y} \quad \mathbf{v}_4 = \begin{pmatrix} -2 \\ -4 \\ 6 \end{pmatrix}$$

Solución Escribimos los vectores como los renglones de una matriz A y reducimos la matriz a la forma escalonada por renglones. La matriz resultante tendrá el mismo espacio de renglones que A.

$$\begin{pmatrix} 1 & 2 & -3 \\ -2 & 0 & 4 \\ 0 & 4 & -2 \\ -2 & -4 & 6 \end{pmatrix} \xrightarrow[A_{1,4}(2)]{A_{1,2}(2)} \begin{pmatrix} 1 & 2 & -3 \\ 0 & 4 & -2 \\ 0 & 4 & -2 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\xrightarrow{A_{2,3}(-1)} \begin{pmatrix} 1 & 2 & -3 \\ 0 & 4 & -2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \xrightarrow{M_2(\frac{1}{4})} \begin{pmatrix} 1 & 2 & -3 \\ 0 & 1 & -\frac{1}{2} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Así pues, una base para gen $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ es $\left\{\begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -\frac{1}{2} \end{pmatrix}\right\}$. Por ejemplo,

$$\begin{pmatrix} -2 \\ 0 \\ 4 \end{pmatrix} = -2\begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix} + 4\begin{pmatrix} 0 \\ 1 \\ -\frac{1}{2} \end{pmatrix}.$$

El siguiente teorema proporciona la relación entre rango y nulidad.

Teorema 5 Sea A una matriz de $m \times n$. Entonces

$$\rho(A) + \nu(A) = n \qquad (5)$$

Demostración Suponemos que $k = \rho(A)$ y que las primeras k columnas de A son linealmente independientes. Sea $\mathbf{c}_i$ $(i > k)$ cualquier otra columna de A. Debido a que $\mathbf{c}_1$,

$\mathbf{c}_2, \ldots, \mathbf{c}_k$ forman una base de C_A, se tiene para algunos escalares $a_1, a_2, \ldots, a_k$

$$\mathbf{c}_i = a_1\mathbf{c}_1 + a_2\mathbf{c}_2 + \cdots + a_k\mathbf{c}_k \tag{6}$$

Así pues, sumando $-a_1\mathbf{c}_1, -a_2\mathbf{c}_2, \ldots, -a_k\mathbf{c}_k$ en forma sucesiva a la i-ésima columna de A, obtenemos una nueva matriz B de $m \times n$ con $\rho(B) = \rho(A)$ y $\nu(B) = \nu(A)$ cuya columna i es $\mathbf{0}$. Así procedemos con todas las demás columnas de A (excepto las primeras k) y resultará la matriz

$$D = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1k} & 0 & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & a_{2k} & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mk} & 0 & 0 & \cdots & 0 \end{pmatrix} \tag{7}$$

donde $\rho(D) = \rho(A)$ y $\nu(D) = \nu(A)$. Con un posible rearreglo de los renglones de D, es posible suponer que los primeros k renglones de D son independientes. Luego se hace lo mismo con los renglones (esto es, sumar múltiplos de las primeras k filas a los últimos $m - k$ renglones) para obtener una nueva matriz:

$$F = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1k} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & a_{2k} & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix}$$

en donde $\rho(F) = \rho(A)$ y $\nu(F) = \nu(A)$. Es obvio ahora que si $i > k$, entonces $F\mathbf{e}_i = \mathbf{0}$,* así que $E_k = \{\mathbf{e}_{k+1}, \mathbf{e}_{k+2}, \ldots, \mathbf{e}_n\}$ es un conjunto linealmente independiente de $n - k$ vectores en N_F. Ahora demostraremos que E_k genera N_F. Sea el vector $\mathbf{x} \in N_F$ con la forma

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \\ \vdots \\ x_n \end{pmatrix}$$

* Recordemos que $\mathbf{e}_i$ es el vector con un valor 1 en la posición i y con valores cero en el resto de las posiciones.

De manera que

$$\mathbf{0} = F\mathbf{x} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1k}x_k \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2k}x_k \\ \vdots \\ a_{k1}x_1 + a_{k2}x_2 + \cdots + a_{kk}x_k \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

El determinante de la matriz del sistema homogéneo de $k \times k$ descrito es distinto de cero, puesto que los renglones de esta matriz son linealmente independientes. Así que la única solución al sistema es $x_1 = x_2 = \cdots = x_k = 0$. Por lo tanto $\mathbf{x}$ posee la forma

$$(0, 0, \ldots, 0, x_{k+1}, x_{k+2}, \ldots, x_n) = x_{k+1}\mathbf{e}_{k+1} + x_{k+2}\mathbf{e}_{k+2} + \cdots + x_n\mathbf{e}_n$$

Lo anterior significa que E_k genera a N_F de modo que $\nu(F) = n - k = n - \rho(F)$. Esto completa la demostración. ■

Ejemplo 7 Para $A = \begin{pmatrix} 1 & 2 & -1 \\ 2 & -1 & 3 \end{pmatrix}$ calculamos (en los Ejemplos 1 y 3) que $\rho(A) = 2$ y que $\nu(A) = 1$; esto ilustra que $\rho(A) + \nu(A) = n(=3)$.

Ejemplo 8 Evaluar $\nu(A)$ para $A = \begin{pmatrix} 1 & -1 & 3 \\ 2 & 0 & 4 \\ -1 & -3 & 1 \end{pmatrix}$

Solución En el Ejemplo 5 encontramos que $\rho(A) = 2$. Así que $\nu(A) = 3 - 2 = 1$.

Teorema 6 Sea A una matriz de $n \times n$. Entonces A es invertible si y sólo si $\rho(A) = n$.

Demostración Por el Teorema 1, A es invertible si y sólo si $\nu(A) = 0$. Por el Teorema 5, $\rho(A) = n - \nu(A)$. Así, A es invertible si y sólo si $\rho(A) = n - 0 = n$. ■

Ahora se mostrará la forma en que la noción de rango puede utilizarse para resolver sistemas de ecuaciones lineales. De nuevo, consideramos el sistema de m ecuaciones en n incógnitas

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \vdots \qquad & \quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned} \tag{8}$$

el cual escribimos como $A\mathbf{x} = \mathbf{b}$. Usamos el símbolo $(A, \mathbf{b})$ para denotar la matriz aumentada $m \times (n + 1)$ obtenida (como en la Sección 1.6) al adjuntar a A el vector $\mathbf{b}$.

Teorema 7 El sistema $A\mathbf{x} = \mathbf{b}$ posee al menos una solución si y sólo si $\mathbf{b} \in C_A$. Esto ocurrirá si y sólo si A y la matriz aumentada $(A, \mathbf{b})$ poseen el mismo rango.

Demostración Si $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n$ son las columnas de A, es posible expresar el sistema (8) como

$$x_1\mathbf{c}_1 + x_2\mathbf{c}_2 + \cdots + x_n\mathbf{c}_n = \mathbf{b} \tag{9}$$

El sistema (9) poseerá una solución (al menos) si y sólo si $\mathbf{b}$ puede escribirse como combinación lineal de las columnas de A. Esto es, para que exista solución, hay que tener $\mathbf{b} \in C_A$. Si $\mathbf{b} \in C_A$, entonces $(A, \mathbf{b})$ posee el mismo número de columnas linealmente independientes que A, de modo que A y $(A, \mathbf{b})$ son el mismo rango. Si $\mathbf{b} \notin C_A$, entonces $\rho(A, \mathbf{b}) = \rho(A) + 1$ y el sistema no tiene solución. Esto completa la demostración. ■

Ejemplo 9 Determinar si el sistema

$$\begin{aligned} 2x_1 + 4x_2 + 6x_3 &= 18 \\ 4x_1 + 5x_2 + 6x_3 &= 24 \\ 2x_1 + 7x_2 + 12x_3 &= 40 \end{aligned}$$

posee soluciones.

Solución Sea $A = \begin{pmatrix} 2 & 4 & 6 \\ 4 & 5 & 6 \\ 2 & 7 & 12 \end{pmatrix}$. Reducimos por filas, para obtener, en forma sucesiva,

$$\xrightarrow{M_1(\frac{1}{2})} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 2 & 7 & 12 \end{pmatrix} \xrightarrow[A_{1,3}(-2)]{A_{1,2}(-4)} \begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & 3 & 6 \end{pmatrix}$$

$$\xrightarrow{M_2(-\frac{1}{3})} \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 3 & 6 \end{pmatrix} \xrightarrow[A_{2,3}(-3)]{A_{2,1}(-2)} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

Por tanto, $\rho(A) = 2$. Similarmente, reducimos $(A, \mathbf{b})$ por filas y obtenemos

$$\left(\begin{array}{ccc|c} 2 & 4 & 6 & 18 \\ 4 & 5 & 6 & 24 \\ 2 & 7 & 12 & 40 \end{array}\right) \xrightarrow{M_1(\frac{1}{2})} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 9 \\ 4 & 5 & 6 & 24 \\ 2 & 7 & 12 & 40 \end{array}\right)$$

$$\xrightarrow[A_{1,3}(-2)]{A_{1,2}(-4)} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 9 \\ 0 & -3 & -6 & -12 \\ 0 & 3 & 6 & 22 \end{array}\right)$$

$$\xrightarrow{M_2(-\frac{1}{3})} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 9 \\ 0 & 1 & 2 & 4 \\ 0 & 3 & 6 & 22 \end{array}\right) \xrightarrow[A_{2,3}(-3)]{A_{2,1}(-2)} \left(\begin{array}{ccc|c} 1 & 0 & -1 & 1 \\ 0 & 1 & 2 & 4 \\ 0 & 0 & 0 & 10 \end{array}\right)$$

Es fácil ver que las últimas tres columnas de la última matriz son linealmente independientes, de modo que $\rho(A, \mathbf{b}) = 3$ y no hay soluciones para el sistema.

Ejemplo 10 Determinar si el sistema

$$\begin{aligned} x_1 - x_2 + 2x_3 &= 4 \\ 2x_1 + x_2 - 3x_3 &= -2 \\ 4x_1 - x_2 + x_3 &= 6 \end{aligned}$$

posee soluciones.

Solución Sea $A = \begin{pmatrix} 1 & -1 & 2 \\ 2 & 1 & -3 \\ 4 & -1 & 1 \end{pmatrix}$. Entonces $\det A = 0$, de manera que $\rho(A) =< 3$.

Como la primera columna no es múltiplo de la segunda, es claro que las primeras dos columnas son linealmente independientes; así pues, $\rho(A) = 2$. Para calcular $\rho(A, \mathbf{b})$, reducimos por filas:

$$\left(\begin{array}{ccc|c} 1 & -1 & 2 & 4 \\ 2 & 1 & -3 & -2 \\ 4 & -1 & 1 & 6 \end{array}\right) \xrightarrow[A_{1,3}(-4)]{A_{1,2}(-2)} \left(\begin{array}{ccc|c} 1 & -1 & 2 & 4 \\ 0 & 3 & -7 & -10 \\ 0 & 3 & -7 & -10 \end{array}\right)$$

Vemos que $\rho(A, \mathbf{b}) = 2$ y existe una infinidad de soluciones del sistema. (Si existiera sólo una, se tendría que $\det A \neq 0$.)

Los resultados de esta sección permiten mejorar nuestro Teorema Resumen — visto por último en la Sección 4.4, página 195.

Teorema 8 ***Teorema Resumen — Versión 6.*** Sea A una matriz de $n \times n$. Entonces son equivalentes los siguientes nueve enunciados. Esto es, si uno es cierto, todos los demás se verifican.

i. A es invertible.
ii. La única solución del sistema homogéneo $A\mathbf{x} = \mathbf{0}$ es la solución trivial $(\mathbf{x} = \mathbf{0})$.
iii. El sistema $A\mathbf{x} = \mathbf{b}$ posee una única solución para todo n-vector $\mathbf{b}$.
iv. A es equivalente por filas a la matriz identidad I_n de $n \times n$.
v. A puede ser expresada como el producto de matrices elementales.
vi. Los renglones (y las columnas) de A son linealmente independientes.
vii. $\det A \neq 0$.
viii. $\nu(A) = 0$.
ix. $\rho(A) = n$.

Más aún, si alguno de los enunciados falla, entonces para cada vector $\mathbf{b} \in \mathbb{R}^n$, el sistema $A\mathbf{x} = \mathbf{b}$ posee una infinidad de soluciones o ninguna. Tendrá una infinidad de soluciones si y sólo si $\rho(A) = \rho((A, \mathbf{b}))$.

Concluiremos esta sección con una demostración del Teorema 3.

Demostración del Teorema 3* Primero hay que probar que C_A = Imag$A$. Como antes, sea $\mathbf{e}_j$ el vector en $\mathbb{R}^n$ con un 1 en la posición j y cero en las demás posiciones. Escribimos A en la forma

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mj} & \cdots & a_{mn} \end{pmatrix}$$

Entonces $A\mathbf{e}_j$ es la j-ésima columna de A. Así pues, cada columna de A está en Imag$A$, de modo que

$$C_A \subseteq \text{Imag}\, A \qquad (10)$$

Sea $\{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_k\}$ una base de Imag A. Examinemos un vector de ellos, el $\mathbf{y}_i$, por ejemplo. Por definición de imagen existe un $\mathbf{x}_i \in \mathbb{R}^n$ tal que $\mathbf{y}_i = A\mathbf{x}_i$. Pero $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ es una base de $\mathbb{R}^n$, así que existen constantes $c_1, c_2, \ldots, c_n$ tales que

$$\mathbf{x}_i = c_1\mathbf{e}_1 + c_2\mathbf{e}_2 + \cdots + c_n\mathbf{e}_n \qquad (11)$$

Entonces,

$$\mathbf{y}_i = A\mathbf{x}_i = A(c_1\mathbf{e}_1 + c_2\mathbf{e}_2 + \cdots + c_n\mathbf{e}_n) = c_1A\mathbf{e}_1 + c_2A\mathbf{e}_2 + \cdots + c_nA\mathbf{e}_n \qquad (12)$$

Pero $A\mathbf{e}_j$ es la j-ésima columna de A, de modo que (12) muestra que $\mathbf{y}_i$ puede expresarse como combinación lineal de las columnas de A. Por consiguiente, cada vector base en Imag$A$ está en el espacio de columnas C_A de A, de forma que

$$\text{Imag}\ A \subseteq C_A \qquad (13)$$

Combinando (10) y (13), vemos que Imag$A$ = C_A. Para completar la demostración, hay que probar que si R_A denota al espacio de renglones de A, entonces $\dim R_A = \dim C_A$. Denotamos los renglones de A por $\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_m$, y sea $k = \dim R_A$. Sea también $S = \{\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_k\}$ una base de R_A. Entonces, todo vector renglón de A puede expresarse como combinación lineal de los vectores en S, y tenemos, para algunas constantes α_{ij}, que:

$$\begin{aligned} \mathbf{r}_1 &= \alpha_{11}\mathbf{s}_1 + \alpha_{12}\mathbf{s}_2 + \cdots + \alpha_{1k}\mathbf{s}_k \\ \mathbf{r}_2 &= \alpha_{21}\mathbf{s}_1 + \alpha_{22}\mathbf{s}_2 + \cdots + \alpha_{2k}\mathbf{s}_k \\ &\vdots \\ \mathbf{r}_m &= \alpha_{m1}\mathbf{s}_1 + \alpha_{m2}\mathbf{s}_2 + \cdots + \alpha_{mk}\mathbf{s}_k \end{aligned} \qquad (14)$$

Ahora, la componente j de $\mathbf{r}_i$ es a_{ij}. Si ahora igualamos las componentes j de ambos miembros de (14),

$$\begin{aligned} a_{1j} &= \alpha_{11}s_{1j} + \alpha_{12}s_{2j} + \cdots + \alpha_{1k}s_{kj} \\ a_{2j} &= \alpha_{21}s_{1j} + \alpha_{22}s_{2j} + \cdots + \alpha_{2k}s_{kj} \\ &\vdots \\ a_{mj} &= \alpha_{m1}s_{1j} + \alpha_{m2}s_{2j} + \cdots + \alpha_{mk}s_{kj} \end{aligned}$$

* Si el tiempo lo permite.

o bien

$$\begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix} = s_{1j}\begin{pmatrix} \alpha_{11} \\ \alpha_{21} \\ \vdots \\ \alpha_{m1} \end{pmatrix} + s_{2j}\begin{pmatrix} \alpha_{12} \\ \alpha_{22} \\ \vdots \\ \alpha_{m2} \end{pmatrix} + \cdots + s_{kj}\begin{pmatrix} \alpha_{1k} \\ \alpha_{2k} \\ \vdots \\ \alpha_{mk} \end{pmatrix} \qquad (15)$$

Aquí $\mathbf{s}_i = (s_{i1}, s_{i2}, \ldots, s_{in})$. Sea $\boldsymbol{\alpha}_i$ el vector $\begin{pmatrix} \alpha_{1i} \\ \alpha_{2i} \\ \vdots \\ \alpha_{mi} \end{pmatrix}$. Entonces, como el primer miembro de (15) es la columna j de A, es posible escribir cada columna de A como combinación lineal de $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_k$, lo cual significa que $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_k$ generan C_A y que

$$\dim C_A \leq k = \dim R_A \qquad (16)$$

Pero la Ecuación (16) se satisface para cualquier matriz A. En particular, se satisface para A^t. Pero $C_{A^t} = R_A$ y $R_{A^t} = C_A$. Por tanto, de (16), $\dim C_{A^t} \leq \dim R_{A^t}$, tenemos

$$\dim R_A \leq \dim C_A. \qquad (17)$$

Combinando (16) y (17) se completa la demostración. ■

Problemas 4.7

En los Problemas del 1 *al* 15, *obtenga el rango y la nulidad de la matriz indicada.*

1. $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$

2. $\begin{pmatrix} 1 & -1 & 2 \\ 3 & 1 & 0 \end{pmatrix}$

3. $\begin{pmatrix} -1 & 3 & 2 \\ 2 & -6 & -4 \end{pmatrix}$

4. $\begin{pmatrix} 1 & -1 & 2 \\ 3 & 1 & 4 \\ -1 & 0 & 4 \end{pmatrix}$

5. $\begin{pmatrix} 1 & -1 & 2 \\ 3 & 1 & 4 \\ 5 & -1 & 8 \end{pmatrix}$

6. $\begin{pmatrix} -1 & 2 & 1 \\ 2 & -4 & -2 \\ -3 & 6 & 3 \end{pmatrix}$

7. $\begin{pmatrix} 1 & -1 & 2 & 3 \\ 0 & 1 & 4 & 3 \\ 1 & 0 & 6 & 6 \end{pmatrix}$

8. $\begin{pmatrix} 1 & -1 & 2 & 3 \\ 0 & 1 & 4 & 3 \\ 1 & 0 & 6 & 5 \end{pmatrix}$

9. $\begin{pmatrix} 2 & 3 \\ -1 & 1 \\ 4 & 7 \end{pmatrix}$

10. $\begin{pmatrix} 1 & -1 & 2 & 3 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$

11. $\begin{pmatrix} 1 & -1 & 2 & 1 \\ -1 & 0 & 1 & 2 \\ 1 & -2 & 5 & 4 \\ 2 & -1 & 1 & -1 \end{pmatrix}$

12. $\begin{pmatrix} 1 & -1 & 2 & 3 \\ -2 & 2 & -4 & -6 \\ 2 & -2 & 4 & 6 \\ 3 & -3 & 6 & 9 \end{pmatrix}$

13. $\begin{pmatrix} -1 & -1 & 0 & 0 \\ 0 & 0 & 2 & 3 \\ 4 & 0 & -2 & 1 \\ 3 & -1 & 0 & 4 \end{pmatrix}$ **14.** $\begin{pmatrix} 3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 6 \end{pmatrix}$ **15.** $\begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & 4 \\ 0 & 0 & 6 \end{pmatrix}$

En los Problemas 16 *a* 22, *obtenga una base para la imagen (o espacio imagen) y el núcleo (o kernel) de la matriz indicada.*

16. La matriz del Problema 2.
17. La matriz del Problema 5.
18. La matriz del Problema 6.
19. La matriz del Problema 8.
20. La matriz del Problema 11.
21. La matriz del Problema 12.
22. La matriz del Problema 13.

En los Problemas del 23 *al* 26 *halle una base para el espacio generado por el conjunto dado de vectores.*

23. $\begin{pmatrix} 1 \\ 4 \\ -2 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} -1 \\ 3 \\ -4 \end{pmatrix}$

24. $(1, -2, 3), (2, -1, 4), (3, -3, 3), (2, 1, 0)$

25. $(1, -1, 1, -1), (2, 0, 0, 1), (4, -2, 2, 1), (7, -3, 3, -1)$

26. $\begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -2 \\ -2 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 2 \\ 1 \end{pmatrix}$

En los Problemas del 27 *al* 30 *utilice el Teorema* 7 *para determinar si el sistema indicado posee soluciones.*

27. $\begin{aligned} x_1 + x_2 - x_3 &= 7 \\ 4x_1 - x_2 + 5x_3 &= 4 \\ 6x_1 + x_2 + 3x_3 &= 20 \end{aligned}$

28. $\begin{aligned} x_1 + x_2 - x_3 &= 7 \\ 4x_1 - x_2 + 5x_3 &= 4 \\ 6x_1 + x_2 + 3x_3 &= 18 \end{aligned}$

29. $\begin{aligned} x_1 - 2x_2 + x_3 + x_4 &= 2 \\ 3x_1 \qquad + 2x_3 - 2x_4 &= -8 \\ 4x_2 - x_3 - x_4 &= 1 \\ 5x_1 \qquad + 3x_3 - x_4 &= -3 \end{aligned}$

30. $\begin{aligned} x_1 - 2x_2 + x_3 + x_4 &= 2 \\ 3x_1 \qquad + 2x_3 - 2x_4 &= -8 \\ 4x_2 - x_3 - x_4 &= 1 \\ 5x_1 \qquad + 3x_3 - x_4 &= 0 \end{aligned}$

31. Demostrar que el rango de una matriz diagonal es igual al número de componentes no nulas en la diagonal.

32. Sea A una matriz de $n \times n$ triangular superior con ceros en la diagonal. Demuestre que $\rho(A) < n$.

33. Demuestre que para cualquier matriz A, $\rho(A) = \rho(A^t)$.

34. Demuestre que si A es una matriz de $m \times n$ y $m < n$, entonces, (**a**) $\rho(A) \le m$ y (**b**) $\nu(A) \ge n - m$.

35. Sea A una matriz de $m \times n$ y sean B y C matrices invertibles de $m \times m$ y de $n \times n$, respectivamente. Demuestre que $\rho(A) = \rho(BA) = \rho(AC)$. Esto es, el multiplicar una matriz por una matriz invertible, no altera su rango.

36. Sean A y B matrices de $m \times n$ y $n \times p$, respectivamente. Demuestre que $\rho(AB) \le$ mín $(\rho(A), \rho(B))$.

37. Sea A una matriz de 5×7 con rango 5. Demuestre que el sistema lineal $A\mathbf{x} = \mathbf{b}$ posee al menos una solución para cada vector-5**b**.

★ **38.** Sean A y B matrices de $m \times n$. Demuestre que si $\rho(A) = \rho(B)$, entonces existen matrices invertibles C y D tales que $B = CAD$.

39. Si $B = CAD$, en donde C y D son invertibles, entonces pruebe que $\rho(A) = \rho(B)$.

40. Suponiendo que cualesquiera k renglones de A son linealmente independientes, mientras que cualesquiera $k + 1$ renglones de A no lo son, pruebe entonces que $\rho(A) = k$.

41. Si A es una matriz de $n \times n$, demuestre que $\rho(A) < n$ si y sólo si existe un vector $\mathbf{x} \in \mathbb{R}^n$ tal que $\mathbf{x} \neq \mathbf{0}$ y $A\mathbf{x} = \mathbf{0}$.

42. Sea A una matriz de $m \times n$. Si para cada $\mathbf{y} \in \mathbb{R}^m$ existe $\mathbf{x} \in \mathbb{R}^n$ tal que $A\mathbf{x} = \mathbf{y}$, demuestre que $\rho(A) = m$.

4.8 Rango y determinantes de submatrices (opcional)

En esta sección mostraremos cómo puede ser evaluado el rango de una matriz mediante el cálculo de ciertos determinantes.

Definición 1 **Submatriz cuadrada de orden k.** Sea A una matriz de $m \times n$. La matriz obtenida como la intersección de k renglones y k columnas de A se llama *submatriz cuadrada de orden k* de A.

Ejemplo 1 Sea

$$A = \begin{pmatrix} 1 & 3 & -2 & 6 \\ 3 & 1 & 6 & 2 \\ 4 & 1 & 5 & 2 \end{pmatrix}$$

i. Cada componente de A es una submatriz cuadrada de orden 1 de A.

ii. La intersección de los renglones 1 y 3 y las columnas 3 y 4 es una submatriz cuadrada de orden 2:

$$\begin{pmatrix} -2 & 6 \\ 5 & 2 \end{pmatrix}$$

Existen dieciocho submatrices de orden 2. Otras dos son

$$\underset{\substack{\text{renglones 1 y 2} \\ \text{columnas 1 y 2}}}{\begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}} \quad \text{y} \quad \underset{\substack{\text{renglones 1 y 3} \\ \text{columnas 2 y 4}}}{\begin{pmatrix} 3 & 6 \\ 1 & 2 \end{pmatrix}}$$

iii. Existen cuatro submatrices cuadradas de orden 3. Éstas son

$$\begin{pmatrix} 1 & 3 & -2 \\ 3 & 1 & 6 \\ 4 & 1 & 5 \end{pmatrix}, \quad \begin{pmatrix} 1 & 3 & 6 \\ 3 & 1 & 2 \\ 4 & 1 & 2 \end{pmatrix}, \quad \begin{pmatrix} 1 & -2 & 6 \\ 3 & 6 & 2 \\ 4 & 5 & 2 \end{pmatrix}, \quad \text{y} \quad \begin{pmatrix} 3 & -2 & 6 \\ 1 & 6 & 2 \\ 1 & 5 & 2 \end{pmatrix}$$

Definición 2 **Subdeterminante de orden k.** Sea A una matriz de $m \times n$. Un *subdeterminante de orden* k de A es el determinante de una submatriz cuadrada de orden k de A.

Ejemplo 2 Sea

$$A = \begin{pmatrix} 1 & 3 & -2 & 6 \\ 3 & 1 & 6 & 2 \\ 4 & 1 & 5 & 2 \end{pmatrix}$$

i. Cada componente de A es un subdeterminante de orden 1.
ii. Tres subdeterminantes de orden 2 son:

$$\begin{vmatrix} -2 & 6 \\ 5 & 2 \end{vmatrix} = -34, \quad \begin{vmatrix} 1 & 3 \\ 3 & 1 \end{vmatrix} = -8, \quad \text{y} \quad \begin{vmatrix} 3 & 6 \\ 1 & 2 \end{vmatrix} = 0$$

iii. Dos subdeterminantes de orden 3 son:

$$\begin{vmatrix} 1 & 3 & -2 \\ 3 & 1 & 6 \\ 4 & 1 & 5 \end{vmatrix} = 28 \quad \text{y} \quad \begin{vmatrix} 3 & -2 & 6 \\ 1 & 6 & 2 \\ 1 & 5 & 2 \end{vmatrix} = 0$$

Teorema 1 Sea A una matriz de $m \times n$. Entonces

> $\rho(A) \geq k$
> si y sólo si A posee un subdeterminante no nulo de orden k.

Demostración Supongamos que $\rho(A) \geq k$. Del Teorema 4.7.3 sabemos que

$$\dim R_A = \dim C_A = \dim \text{Imag } A = \rho(A) \geq k$$

Así pues, A posee al menos k renglones independientes. Sea B una matriz de $k \times n$ cuyos renglones son k renglones de A linealmente independientes. Entonces $k = \rho(B) = \dim C_B$; esto es, B tiene k columnas linealmente independientes. Si E es la matriz de $k \times k$ cuyas columnas son k columnas linealmente independientes de B, entonces E es una submatriz de A de orden k. Como las columnas de E son linealmente independientes, entonces $\det E \neq 0$. Hemos demostrado que A posee un subdeterminante de orden k distinto de cero.

Supóngase ahora que A posee un subdeterminante no nulo de orden k; esto es, que A tiene una submatriz cuadrada E de orden k tal que $\det E \neq 0$. Entonces los renglones de E son linealmente independientes. Cada uno de los renglones de E está contenido en un renglón de A. Así, los renglones correspondientes de A son linealmente independientes. Como A tiene al menos k renglones linealmente independientes, concluimos que $\rho(A) \geq k$. ■

Teorema 2 Sea A una matriz de $m \times n$, entonces,

> $\rho(A) = r > 0$ si y sólo si A posee al menos un subdeterminante no nulo de orden r y cada subdeterminante de orden $r + 1$ es cero.

Demostración Supongamos que $\rho(A) = r$. Entonces, por el Teorema 1, A posee un subdeterminante no nulo de orden r. Sea E una submatriz cuadrada de A de orden $(r + 1)$. Así, $\det E = 0$, pues entonces tendríamos que $\rho(A) \geq k + 1$ por el Teorema 1. Recíprocamente, supóngase que A posee una submatriz cuadrada E, de orden r, tal que $\det E \neq 0$ y que además es cero todo subdeterminante de orden $r + 1$. Por el Teorema 1 de nuevo, $\rho(A) \geq r$, pero $\rho(A) < r + 1$ pues si suponemos que $\rho(A) \geq r + 1$, entonces A posee un subdeterminante no nulo de orden $r + 1$, en contradicción con la hipótesis. Así pues, $\rho(A) = k$, y el teorema queda demostrado. ■

Ejemplo 3 Usar el Teorema 2 para encontrar el rango de cada matriz.

$$\textbf{(a)}\ A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ -3 & -6 & -9 \end{pmatrix} \qquad \textbf{(b)}\ A = \begin{pmatrix} 1 & 2 & -1 & 4 \\ 3 & 1 & 5 & 2 \\ 1 & -3 & 7 & -6 \end{pmatrix}$$

Solución **(a)** A posee nueve subdeterminantes de orden 1. Los nueve subdeterminantes de A de orden 2 son:

$$\begin{vmatrix} 1 & 2 \\ 2 & 4 \end{vmatrix} = \begin{vmatrix} 1 & 3 \\ 2 & 6 \end{vmatrix} = \begin{vmatrix} 2 & 3 \\ 4 & 6 \end{vmatrix} = \begin{vmatrix} 1 & 2 \\ -3 & -6 \end{vmatrix} = \begin{vmatrix} 1 & 3 \\ -3 & -9 \end{vmatrix} = \begin{vmatrix} 2 & 3 \\ -6 & -9 \end{vmatrix}$$

$$= \begin{vmatrix} 2 & 4 \\ -3 & -6 \end{vmatrix} = \begin{vmatrix} 2 & 6 \\ -3 & -9 \end{vmatrix} = \begin{vmatrix} 4 & 6 \\ -6 & -9 \end{vmatrix} = 0$$

Así pues $\rho(A) = 1$.

(b) $\begin{vmatrix} 1 & 2 \\ 3 & 1 \end{vmatrix} = -5 \neq 0$. De modo que $\rho(A) \geq 2$. Pero

$$\begin{vmatrix} 1 & 2 & -1 \\ 3 & 1 & 5 \\ 1 & -3 & 7 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 4 \\ 3 & 1 & 2 \\ 1 & -3 & -6 \end{vmatrix} = \begin{vmatrix} 2 & -1 & 4 \\ 1 & 5 & 2 \\ -3 & 7 & -6 \end{vmatrix} = 0$$

Verificar esto

En consecuencia, $\rho(A) = 2$.

Problemas 4.8

En los Problemas del 1 *al* 10, *utilizar el Teorema* 2 *para determinar el rango de la matriz indicada.*

1. $\begin{pmatrix} 2 & -3 \\ -6 & 9 \end{pmatrix}$ **2.** $\begin{pmatrix} 4 & 1 \\ 1 & 4 \end{pmatrix}$ **3.** $\begin{pmatrix} 1 & 0 & 6 \\ 2 & -1 & 1 \end{pmatrix}$ **4.** $\begin{pmatrix} 1 & 0 & 2 \\ 3 & 0 & 6 \end{pmatrix}$

5. $\begin{pmatrix} 1 & 0 & 2 \\ 3 & 1 & 4 \\ 1 & 1 & 0 \end{pmatrix}$ **6.** $\begin{pmatrix} 1 & 0 & 2 \\ 3 & 1 & 4 \\ 1 & 1 & 1 \end{pmatrix}$ **7.** $\begin{pmatrix} 1 & 3 & -1 & 4 \\ 2 & 1 & 4 & 1 \\ 1 & 0 & 2 & 3 \end{pmatrix}$

8. $\begin{pmatrix} 1 & 3 & -1 & 4 \\ 2 & 1 & 4 & 1 \\ -1 & 2 & -5 & 3 \end{pmatrix}$ **9.** $\begin{pmatrix} 1 & 2 & 1 \\ 2 & -1 & 4 \\ 4 & 3 & 6 \\ 2 & 5 & 2 \end{pmatrix}$ **10.** $\begin{pmatrix} 1 & 2 & 1 \\ 2 & -1 & 4 \\ 4 & 3 & 6 \\ 7 & 4 & 4 \end{pmatrix}$.

11. Sea A una matriz de $m \times n$ y supóngase que todo subdeterminante de orden k es cero. Demuestre que todo subdeterminante de orden $k + 1$ es cero.

4.9 Cambio de base

En $\mathbb{R}^2$ se formuló cualquier vector en términos de la base *estándar* (o *canónica*) $\mathbf{i} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\mathbf{j} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. En $\mathbb{R}^n$ ya definimos la base estándar $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$, y en P_n, la base estándar con $\{1, x, x^2, \ldots, x^n\}$. Estas bases son las más usadas puesto que es muy fácil trabajar con ellas. Sin embargo, en ciertas ocasiones son convenientes otras bases. Obviamente, existen muchas bases que podemos escoger puesto que en un espacio vectorial de dimensión n, *cualquier* conjunto de n vectores linealmente independientes forma una base. En esta sección veremos cómo cambiar de una base a otra calculando cierta matriz.

Empezamos con un ejemplo muy simple. Sean $\mathbf{u}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ y $\mathbf{u}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Entonces $B_1 = \{\mathbf{u}_1, \mathbf{u}_2\}$ es la base estándar en $\mathbb{R}^2$. Sean $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$ y $\mathbf{v}_2 = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$. Como $\mathbf{v}_1$ y $\mathbf{v}_2$ son linealmente independientes (puesto que $\mathbf{v}_1$ no es un múltiplo de $\mathbf{v}_2$), $B_2 = \{\mathbf{v}_1, \mathbf{v}_2\}$ es una segunda base en $\mathbb{R}^2$. Sea $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ un vector en $\mathbb{R}^2$. Esta notación significa que

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1\begin{pmatrix} 1 \\ 0 \end{pmatrix} + x_2\begin{pmatrix} 0 \\ 1 \end{pmatrix} = x_1\mathbf{u}_1 + x_2\mathbf{u}_2.$$

Esto es, $\mathbf{x}$ se exprese en términos de los vectores en la base B_1. Para destacar esto, se escribe

$$(\mathbf{x})_{B_1} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Como B_2 es otra base de $\mathbb{R}^2$, existen escalares c_1 y c_2 tales que

$$\mathbf{x} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 \tag{1}$$

Una vez encontrados estos escalares, se tiene

$$(\mathbf{x})_{B_2} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$$

para indicar que ahora $\mathbf{x}$ está expresado en términos de los vectores en B_2. Para encontrar los números c_1 y c_2, escribimos la base anterior ($\mathbf{u}_1$ y $\mathbf{u}_2$) en términos de los vectores de la nueva base ($\mathbf{v}_1$ y $\mathbf{v}_2$). Es fácil verificar que

$$\mathbf{u}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \tfrac{2}{5}\begin{pmatrix} 1 \\ 3 \end{pmatrix} - \tfrac{3}{5}\begin{pmatrix} -1 \\ 2 \end{pmatrix} = \tfrac{2}{5}\mathbf{v}_1 - \tfrac{3}{5}\mathbf{v}_2 \tag{2}$$

y

$$\mathbf{u}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \tfrac{1}{5}\begin{pmatrix} 1 \\ 3 \end{pmatrix} + \tfrac{1}{5}\begin{pmatrix} -1 \\ 2 \end{pmatrix} = \tfrac{1}{5}\mathbf{v}_1 + \tfrac{1}{5}\mathbf{v}_2 \tag{3}$$

Esto es,

$$(\mathbf{u}_1)_{B_2} = \begin{pmatrix} \frac{2}{5} \\ -\frac{3}{5} \end{pmatrix} \quad \text{y} \quad (\mathbf{u}_2)_{B_2} = \begin{pmatrix} \frac{1}{5} \\ \frac{1}{5} \end{pmatrix}$$

Entonces

$$\mathbf{x} = x_1\mathbf{u}_1 + x_2\mathbf{u}_2 \overset{\text{de (2) y (3)}}{=} x_1(\tfrac{2}{5}\mathbf{v}_1 - \tfrac{3}{5}\mathbf{v}_2) + x_2(\tfrac{1}{5}\mathbf{v}_1 + \tfrac{1}{5}\mathbf{v}_2)$$
$$= (\tfrac{2}{5}x_1 + \tfrac{1}{5}x_2)\mathbf{v}_1 + (-\tfrac{3}{5}x_1 + \tfrac{1}{5}x_2)\mathbf{v}_2$$

Por consiguiente, de (1),

$$c_1 = \tfrac{2}{5}x_1 + \tfrac{1}{5}x_2$$
$$c_2 = -\tfrac{3}{5}x_1 + \tfrac{1}{5}x_2$$

o bien

$$(\mathbf{x})_{B_2} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} \frac{2}{5}x_1 + \frac{1}{5}x_2 \\ -\frac{3}{5}x_1 + \frac{1}{5}x_2 \end{pmatrix} = \begin{pmatrix} \frac{2}{5} & \frac{1}{5} \\ -\frac{3}{5} & \frac{1}{5} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Por ejemplo, si $(\mathbf{x})_{B_1} = \begin{pmatrix} 3 \\ -4 \end{pmatrix}$, entonces

$$(\mathbf{x})_{B_2} = \begin{pmatrix} \frac{2}{5} & \frac{1}{5} \\ -\frac{3}{5} & \frac{1}{5} \end{pmatrix}\begin{pmatrix} 3 \\ -4 \end{pmatrix} = \begin{pmatrix} \frac{2}{5} \\ -\frac{13}{5} \end{pmatrix}$$

Verificación

$$\tfrac{2}{5}\mathbf{v}_1 - \tfrac{13}{5}\mathbf{v}_2 = \tfrac{2}{5}\begin{pmatrix} 1 \\ 3 \end{pmatrix} - \tfrac{13}{5}\begin{pmatrix} -1 \\ 2 \end{pmatrix} = \begin{pmatrix} \frac{2}{5} + \frac{13}{5} \\ \frac{6}{5} - \frac{26}{5} \end{pmatrix} = \begin{pmatrix} 3 \\ -4 \end{pmatrix} = 3\begin{pmatrix} 1 \\ 0 \end{pmatrix} - 4\begin{pmatrix} 0 \\ 1 \end{pmatrix}$$
$$= 3\mathbf{u}_1 - 4\mathbf{u}_2.$$

La matriz $A = \begin{pmatrix} \frac{2}{5} & \frac{1}{5} \\ -\frac{3}{5} & \frac{1}{5} \end{pmatrix}$ se llama *matriz de transición* de B_1 a B_2, y hemos demostrado que

$$(\mathbf{x})_{B_2} = A(\mathbf{x})_{B_1} \tag{4}$$

Se puede generalizar fácilmente este ejemplo, pero primero necesitamos ampliar la notación. Sean $B_1 = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ y $B_2 = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ dos bases para un espacio vectorial real V de n dimensiones. Sea $\mathbf{x} \in V$. Entonces $\mathbf{x}$ se puede expresar en términos de ambas bases:

$$\mathbf{x} = b_1\mathbf{u}_1 + b_2\mathbf{u}_2 + \cdots + b_n\mathbf{u}_n \tag{5}$$

y

$$\mathbf{x} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n \tag{6}$$

en donde las b_i y las c_i son números reales. Entonces escribimos $(\mathbf{x})_{B_1} =$

$\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$ para denotar la representación de **x** en términos de la base $\mathbf{B}_1$. Esto no se presta a confusión porque los coeficientes b_i en (5) son únicos por el Teorema 4.6.1. Asimismo, $(\mathbf{x})_{B_2} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$ tiene un significado semejante.

Supóngase que $\mathbf{w}_1 = a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + \cdots + a_n\mathbf{u}_n$ y $\mathbf{w}_2 = b_1\mathbf{u}_1 + b_2\mathbf{u}_2 + \cdots + b_n\mathbf{u}_n$. Entonces $\mathbf{w}_1 + \mathbf{w}_2 + (a_1 + b_1)\mathbf{u}_1 + (a_2 + b_2)\mathbf{u}_2 + \cdots + (a_n + b_n)\mathbf{u}_n$, de modo que

$$(\mathbf{w}_1 + \mathbf{w}_2)_{B_1} = (\mathbf{w}_1)_{B_1} + (\mathbf{w}_2)_{B_1}$$

Es decir, con la nueva notación se pueden sumar vectores como se suman en $\mathbb{R}^n$. Además, resulta sencillo demostrar que

$$\alpha(\mathbf{w})_{B_1} = (\alpha\mathbf{w})_{B_1}$$

Ahora bien, como B_2 es una base, cada $\mathbf{u}_j$ en B_1 se puede escribir como una combinación lineal de las $\mathbf{v}_i$. Por lo tanto, existe un sólo conjunto de escalares $a_{1j}, a_{2j}, \ldots, a_{nj}$ tal que para $j = 1, 2, \ldots, n$

$$\mathbf{u}_j = a_{1j}\mathbf{v}_1 + a_{2j}\mathbf{v}_2 + \cdots + a_{nj}\mathbf{v}_n \tag{7}$$

o bien

$$(\mathbf{u}_j)_{B_2} = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{pmatrix} \tag{8}$$

Definición 1 **Matriz de transición.** La matriz A de $n \times n$, cuyas columnas están dadas por (8), recibe el nombre de *matriz de transición* de la base B_1 a la base B_2. Esto es,

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{pmatrix} \tag{9}$$
$$\quad \uparrow \qquad \uparrow \qquad \uparrow \qquad\qquad \uparrow$$
$$(\mathbf{u}_1)_{B_2} \ (\mathbf{u}_2)_{B_2} \ (\mathbf{u}_3)_{B_2} \ \cdots \ (\mathbf{u}_n)_{B_2}$$

Teorema 1 Sean B_1 y B_2 bases para un espacio vectorial V. Sea A la matriz de transición de B_1 a B_2. Entonces, para todo $\mathbf{x} \in V$,

$$(\mathbf{x})_{B_2} = A(\mathbf{x})_{B_1} \tag{10}$$

Demostración Utilizamos la representación de **x** dada en (5) y (6):

$$\mathbf{x} \overset{\text{de (5)}}{=} b_1\mathbf{u}_1 + b_2\mathbf{u}_2 + \cdots + b_n\mathbf{u}_n$$

$$\overset{\text{de (7)}}{=} b_1(a_{11}\mathbf{v}_1 + a_{21}\mathbf{v}_2 + \cdots + a_{n1}\mathbf{v}_n) + b_2(a_{12}\mathbf{v}_1 + a_{22}\mathbf{v}_2 + \cdots + a_{n2}\mathbf{v}_n) + \cdots + b_n(a_{1n}\mathbf{v}_1 + a_{2n}\mathbf{v}_2 + \cdots + a_{nn}\mathbf{v}_n)$$

$$= (a_{11}b_1 + a_{12}b_2 + \cdots + a_{1n}b_n)\mathbf{v}_1 + (a_{21}b_1 + a_{22}b_2 + \cdots + a_{2n}b_n)\mathbf{v}_2 + \cdots + (a_{n1}b_1 + a_{n2}b_2 + \cdots + a_{nn}b_n)\mathbf{v}_n$$

$$\overset{\text{de (6)}}{=} c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n \tag{11}$$

En consecuencia,

$$(\mathbf{x})_{B_2} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} \overset{\text{de (11)}}{=} \begin{pmatrix} a_{11}b_1 + a_{12}b_2 + \cdots + a_{1n}b_n \\ a_{21}b_1 + a_{22}b_2 + \cdots + a_{2n}b_n \\ \vdots \\ a_{n1}b_1 + a_{n2}b_2 + \cdots + a_{nn}b_n \end{pmatrix}$$

$$= \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = A(\mathbf{x})_{B_1} \ ■ \tag{12}$$

Antes de presentar más ejemplos, demostraremos un teorema que es muy útil para los cálculos.

Teorema 2 Si A es la matriz de transición de B_1 a B_2, entonces A^{-1} es la matriz de transición de B_2 a B_1.

Demostración Sea C la matriz de transición de B_2 a B_1. Entonces, de (10), se tiene que

$$(\mathbf{x})_{B_1} = C(\mathbf{x})_{B_2} \tag{13}$$

Pero $(\mathbf{x})_{B_2} = a(\mathbf{x})_{B_1}$, y sustituyendo lo anterior en (13), resulta

$$(\mathbf{x})_{B_1} = CA(\mathbf{x})_{B_1} \tag{14}$$

Se deja como ejercicio (vea el Problema 39) demostrar que (14) puede ser válida para todo **x** en V sólo si $CA = I$. Por consiguiente, del Teorema 1.8.8, $C = A^{-1}$, y el teorema queda demostrado. ■

Observación. Este teorema facilita en especial hallar la matriz de transición de la base estándar $B_1 = \{\mathbf{e}_1, e_2, \ldots, \mathbf{e}_n\}$ en $\mathbb{R}^n$ a cualquier otra base en $\mathbb{R}^n$. Sea $B_2 = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ cualquier otra base y sea C la matriz cuyas columnas son los vectores $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$. Entonces C es la matriz de transición de B_2 a B_1

puesto que cada vector $\mathbf{v}_i$ ya está expresado en términos de la base canónica o estándar. Por ejemplo,

$$\begin{pmatrix}1\\3\\-2\\4\end{pmatrix}_{B_1} = \begin{pmatrix}1\\3\\-2\\4\end{pmatrix} = 1\begin{pmatrix}1\\0\\0\\0\end{pmatrix} + 3\begin{pmatrix}0\\1\\0\\0\end{pmatrix} - 2\begin{pmatrix}0\\0\\1\\0\end{pmatrix} + 4\begin{pmatrix}0\\0\\0\\1\end{pmatrix}$$

Por lo tanto, la matriz de transición de B_1 a B_2 es C^{-1}.

Procedimiento para encontrar la matriz de transición de una base estándar a una base $B_2 = \{\mathbf{v}_1, \mathbf{v}_2, \ldots \mathbf{v}_n\}$

i. Escribir la matriz C cuyas columnas son $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$.
ii. Calcular C^{-1}. Ésta es la matriz de transición requerida.

Ejemplo 1 En $\mathbb{R}^3$ sean $B_1 = \{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ y $B_2 = \left\{\begin{pmatrix}1\\0\\2\end{pmatrix}, \begin{pmatrix}3\\-1\\0\end{pmatrix}, \begin{pmatrix}0\\1\\-2\end{pmatrix}\right\}$. Si $\mathbf{x} = \begin{pmatrix}x\\y\\z\end{pmatrix} \in \mathbb{R}^3$, escriba $\mathbf{x}$ en términos de los vectores en B_2.

Solución Verifiquemos que B_2 es una base. Esto es evidente ya que $\begin{vmatrix}1 & 3 & 0\\0 & -1 & 1\\2 & 0 & -2\end{vmatrix} = 8 \neq 0$.

Dado que $\mathbf{u}_1 = \begin{pmatrix}1\\0\\0\end{pmatrix}$, $\mathbf{u}_2 = \begin{pmatrix}0\\1\\0\end{pmatrix}$, y $\mathbf{u}_3 = \begin{pmatrix}0\\0\\1\end{pmatrix}$, vemos inmediatamente que la matriz de transición, C, de B_2 a B_1 está dada por

$$C = \begin{pmatrix}1 & 3 & 0\\0 & -1 & 1\\2 & 0 & -2\end{pmatrix}$$

Consecuentemente, del Teorema 2, la matriz de transición A de B_1 a B_2 es

$$A = C^{-1} = \tfrac{1}{8}\begin{pmatrix}2 & 6 & 3\\2 & -2 & -1\\2 & 6 & -1\end{pmatrix}$$

Por ejemplo, si $(\mathbf{x})_{B_1} = \begin{pmatrix}1\\-2\\4\end{pmatrix}$, entonces

$$(\mathbf{x})_{B_2} = \tfrac{1}{8}\begin{pmatrix}2 & 6 & 3\\2 & -2 & -1\\2 & 6 & -1\end{pmatrix}\begin{pmatrix}1\\-2\\4\end{pmatrix} = \tfrac{1}{8}\begin{pmatrix}2\\2\\-14\end{pmatrix} = \begin{pmatrix}\frac{1}{4}\\\frac{1}{4}\\-\frac{7}{4}\end{pmatrix}$$

Como comprobación, nótese que

$$\tfrac{1}{4}\begin{pmatrix}1\\0\\2\end{pmatrix} + \tfrac{1}{4}\begin{pmatrix}3\\-1\\0\end{pmatrix} - \tfrac{7}{4}\begin{pmatrix}0\\1\\-2\end{pmatrix} = \begin{pmatrix}1\\-2\\4\end{pmatrix} = 1\begin{pmatrix}1\\0\\0\end{pmatrix} - 2\begin{pmatrix}0\\1\\0\end{pmatrix} + 4\begin{pmatrix}0\\0\\1\end{pmatrix}$$

Ejemplo 2 En P_2 la base estándar es $B_1 = \{1, x, x^2\}$. En consecuencia, otra base es $B_2 = \{4x - 1, 2x^2 - x, 3x^2 + 3\}$. Si $p = a_0 + a_1x + a_2x^2$, escriba p en términos de los polinomios en B_2.

Solución Verifiquemos primero que B_2 es una base. Si $c_1(4x - 1) + c_2(2x^2 - x) + c_3(3x^2 + 3) = 0$ para todo x, entonces, ordenando los términos obtenemos

$$(-c_1 + 3c_3)1 + (4c_1 - c_2)x + (2c_2 + 3c_3)x^2 = 0$$

Pero, al ser $\{1, x, x^2\}$ un conjunto linealmente independiente, debemos tener que

$$\begin{aligned}-c_1 \quad\;\; + 3c_3 &= 0\\ 4c_1 - c_2 \qquad &= 0\\ 2c_2 + 3c_3 &= 0\end{aligned}$$

El determinante de este sistema homogéneo es $\begin{vmatrix}-1 & 0 & 3\\4 & -1 & 0\\0 & 2 & 3\end{vmatrix} = 27 \neq 0$, lo cual significa que $c_1 = c_2 = c_3 = 0$ es la única solución. Además $(4x - 1)_{B_1} = \begin{pmatrix}-1\\4\\0\end{pmatrix}$, $(2x^2 - x)_{B_1} = \begin{pmatrix}0\\-1\\2\end{pmatrix}$, y $(3 + 3x^2)_{B_1} = \begin{pmatrix}3\\0\\3\end{pmatrix}$. Por lo tanto,

$$C = \begin{pmatrix}-1 & 0 & 3\\4 & -1 & 0\\0 & 2 & 3\end{pmatrix}$$

es la matriz de B_2 a B_1, de modo que

$$A = C^{-1} = \tfrac{1}{27}\begin{pmatrix}-3 & 6 & 3\\-12 & -3 & 12\\8 & 2 & 1\end{pmatrix}$$

es la matriz de transición de B_1 a B_2. Como $(a_0 + a_1x + a_2x^2)_{B_1} = \begin{pmatrix}a_0\\a_1\\a_2\end{pmatrix}$, se tiene que

$$(a_0 + a_1x + a_2x^2)_{B_2} = \tfrac{1}{27}\begin{pmatrix}-3 & 6 & 3\\-12 & -3 & 12\\8 & 2 & 1\end{pmatrix}\begin{pmatrix}a_0\\a_1\\a_2\end{pmatrix}$$

$$= \begin{pmatrix}\frac{1}{27}[-3a_0 + 6a_1 + 3a_2]\\\frac{1}{27}[-12a_0 - 3a_1 + 12a_2]\\\frac{1}{27}[8a_0 + 2a_1 + a_2]\end{pmatrix}$$

Por ejemplo, si $p(x) = 5x^2 - 3x + 4$, entonces

$$(5x^2 - 3x + 4)_{B_2} = \frac{1}{27}\begin{pmatrix} -3 & 6 & 3 \\ -12 & -3 & 12 \\ 8 & 2 & 1 \end{pmatrix}\begin{pmatrix} 4 \\ -3 \\ 5 \end{pmatrix} = \begin{pmatrix} -\frac{15}{27} \\ \frac{21}{27} \\ \frac{31}{27} \end{pmatrix}$$

o bien

verificar esto

$$5x^2 - 3x + 4 = -\tfrac{15}{27}(4x - 1) + \tfrac{21}{27}(2x^2 - x) + \tfrac{31}{27}(3x^2 + 3)$$

Ejemplo 3 Sean $B_1 = \left\{\begin{pmatrix} 3 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ -1 \end{pmatrix}\right\}$ y $B_2 = \left\{\begin{pmatrix} 2 \\ 4 \end{pmatrix}, \begin{pmatrix} -5 \\ 3 \end{pmatrix}\right\}$ dos bases en $\mathbb{R}^2$. Si $(\mathbf{x})_{B_1} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$, escribir $\mathbf{x}$ en términos de los vectores en B_2.

Solución Este problema es un poco más difícil porque ninguna de las bases es la base canónica. Hay que escribir los vectores de B_1 como combinaciones lineales de los vectores de B_2. Esto es, se tienen que hallar constantes a_{11}, a_{21}, a_{12}, a_{22}, tales que

$$\begin{pmatrix} 3 \\ 1 \end{pmatrix} = a_{11}\begin{pmatrix} 2 \\ 4 \end{pmatrix} + a_{21}\begin{pmatrix} -5 \\ 3 \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} 2 \\ -1 \end{pmatrix} = a_{12}\begin{pmatrix} 2 \\ 4 \end{pmatrix} + a_{22}\begin{pmatrix} -5 \\ 3 \end{pmatrix}$$

Esto lleva a los sistemas siguientes:

$$\begin{aligned} 2a_{11} - 5a_{21} &= 3 \\ 4a_{11} + 3a_{21} &= 1 \end{aligned} \quad \text{y} \quad \begin{aligned} 2a_{12} - 5a_{22} &= 2 \\ 4a_{12} + 3a_{22} &= -1 \end{aligned}$$

Las soluciones son $a_{11} = \frac{7}{13}$, $a_{21} = -\frac{5}{13}$, $a_{12} = \frac{1}{26}$, y $a_{22} = -\frac{5}{13}$. Por tanto,

$$A = \frac{1}{26}\begin{pmatrix} 14 & 1 \\ -10 & -10 \end{pmatrix}$$

y

$$(\mathbf{x})_{B_2} = \frac{1}{26}\begin{pmatrix} 14 & 1 \\ -10 & -10 \end{pmatrix}\begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} \frac{1}{26}(14b_1 + b_2) \\ -\frac{10}{26}(b_1 + b_2) \end{pmatrix}$$

Si por ejemplo, sea $\mathbf{x} \overset{\text{en base canónica}}{=} \begin{pmatrix} 7 \\ 4 \end{pmatrix}$. Entonces

$$\begin{pmatrix} 7 \\ 4 \end{pmatrix}_{B_1} = b_1\begin{pmatrix} 3 \\ 1 \end{pmatrix} + b_2\begin{pmatrix} 2 \\ -1 \end{pmatrix} = 3\begin{pmatrix} 3 \\ 1 \end{pmatrix} - \begin{pmatrix} 2 \\ -1 \end{pmatrix}$$

de donde,

$$\begin{pmatrix} 7 \\ 4 \end{pmatrix}_{B_1} = \begin{pmatrix} 3 \\ -1 \end{pmatrix}$$

y

$$\begin{pmatrix} 7 \\ 4 \end{pmatrix}_{B_2} = \frac{1}{26}\begin{pmatrix} 14 & 1 \\ -10 & -10 \end{pmatrix}\begin{pmatrix} 3 \\ -1 \end{pmatrix} = \begin{pmatrix} \frac{41}{26} \\ -\frac{20}{26} \end{pmatrix}$$

Esto es

$$\begin{pmatrix} 7 \\ 4 \end{pmatrix} \overset{\text{¡comprobar!}}{=} \frac{41}{26}\begin{pmatrix} 2 \\ 4 \end{pmatrix} - \frac{20}{26}\begin{pmatrix} -5 \\ 3 \end{pmatrix}$$

Usando la notación de esta sección podemos obtener una forma simple de determinar si un conjunto de vectores en un espacio vectorial de dimensión finita es linealmente dependiente o independiente.

Teorema 3 Sea $B_1 = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ una base para el espacio vectorial V de dimensión n. Supongamos que

$$(\mathbf{x}_1)_{B_1} = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix}, (\mathbf{x}_2)_{B_1} = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{pmatrix}, \ldots, (\mathbf{x}_n)_{B_1} = \begin{pmatrix} a_{1n} \\ a_{21} \\ \vdots \\ a_{nn} \end{pmatrix}$$

Sea

$$A = \begin{pmatrix} a_{11} & a_{12} \cdots & a_{1n} \\ a_{21} & a_{22} \cdots & a_{2n} \\ \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} \cdots & a_{nn} \end{pmatrix}$$

entonces $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ son linealmente independientes si y sólo si $\det A \neq 0$.

Demostración Denotemos con $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ las columnas de A. Supongamos que

$$c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots + c_n\mathbf{x}_n = \mathbf{0} \tag{15}$$

Entonces, al usar la suma definida anteriormente en esta sección, se puede escribir (15) como

$$(c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + \cdots + c_n\mathbf{a}_n)_{B_1} = (\mathbf{0})_{B_1} \tag{16}$$

La Ecuación (16) da dos representaciones del vector cero en V en términos de los vectores base en B_1, y ya que la representación en términos de los vectores base es única (Teorema 4.6.1) concluimos que

$$c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + \cdots + c_n\mathbf{a}_n = \mathbf{0} \tag{17}$$

en donde el cero del lado derecho es el vector cero en $\mathbb{R}^n$. Esto demuestra el teorema, puesto que la Ecuación (17) involucra las columnas de A que son linealmente independientes si y sólo si el $\det A \neq 0$. ■

Ejemplo 4 En P_2, determine si los polinomios $3 - x$, $2 + x^2$ y $4 + 5x - 2x^2$ son linealmente independientes o dependientes.

Solución Usando la base $B_1 = \{1, x, x^2\}$, tenemos que $(3 - x)_{B_1} = \begin{pmatrix} 3 \\ -1 \\ 0 \end{pmatrix}$, $(2 + x^2)_{B_1} = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$ y $(4 + 5x - 2x^2)_{B_1} = \begin{pmatrix} 4 \\ 5 \\ -2 \end{pmatrix}$. Entonces, $\det A = \begin{vmatrix} 3 & 2 & 4 \\ -1 & 0 & 5 \\ 0 & 1 & -2 \end{vmatrix} = -23 \neq 0$, de donde concluimos que los polinomios dados son independientes.

Ejemplo 5 En M_{22}, determine si las matrices $\begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix}$, $\begin{pmatrix} -1 & 3 \\ -1 & 1 \end{pmatrix}$, $\begin{pmatrix} 2 & -1 \\ 0 & 1 \end{pmatrix}$, y $\begin{pmatrix} 1 & 4 \\ 4 & 9 \end{pmatrix}$ son linealmente independientes o dependientes.

Solución Usando la base estándar $B_1 = \left\{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right\}$ se obtiene

$$\det A = \begin{vmatrix} 1 & -1 & 2 & 1 \\ 2 & 3 & -1 & 4 \\ 3 & -1 & 0 & 4 \\ 6 & 1 & 1 & 9 \end{vmatrix} = 0$$

de tal forma que las matrices son dependientes. Note que $\det A = 0$, porque el cuarto renglón de A es la suma de los primeros tres renglones. Observemos también que

$$-29\begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix} - 7\begin{pmatrix} -1 & 3 \\ -1 & 1 \end{pmatrix} + \begin{pmatrix} 2 & -1 \\ 0 & 1 \end{pmatrix} + 20\begin{pmatrix} 1 & 4 \\ 4 & 9 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

que muestra la dependencia lineal de las cuatro matrices.

Problemas 4.9

En los Problemas del 1 al 5 escriba $\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$ en términos de la base dada.

1. $\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}$
2. $\begin{pmatrix} 2 \\ -3 \end{pmatrix}, \begin{pmatrix} 3 \\ -2 \end{pmatrix}$
3. $\begin{pmatrix} 5 \\ 7 \end{pmatrix}, \begin{pmatrix} 3 \\ -4 \end{pmatrix}$
4. $\begin{pmatrix} -1 \\ -2 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \end{pmatrix}$
5. $\begin{pmatrix} a \\ c \end{pmatrix}, \begin{pmatrix} b \\ d \end{pmatrix}$, en donde $ad - bc \neq 0$.

En los Problemas del 6 al 10 escriba $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3$ en términos de la base dada.

6. $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$
7. $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$
8. $\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$

9. $\begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}, \begin{pmatrix} -1 \\ 4 \\ 5 \end{pmatrix}, \begin{pmatrix} 3 \\ -2 \\ -4 \end{pmatrix}$
10. $\begin{pmatrix} a \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} b \\ d \\ 0 \end{pmatrix}, \begin{pmatrix} c \\ e \\ f \end{pmatrix}$, en donde $adf \neq 0$.

En los Problemas del 11 al 13 escriba el polinomio $a_0 + a_1x + a_2x^2$ en P_2 en términos de la base dada.

11. $1, x - 1, x^2 - 1$
12. $6, 2 + 3x, 3 + 4x + 5x^2$
13. $x + 1, x - 1, x^2 - 1$
14. En M_{22} escriba la matriz $\begin{pmatrix} 2 & -1 \\ 4 & 6 \end{pmatrix}$ en términos de la base $\left\{\begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 2 & 0 \\ 3 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -2 \\ 0 & 4 \end{pmatrix}\right\}$.
15. En P_3 escriba el polinomio $2x^3 - 3x^2 + 5x - 6$ en términos de los polinomios base 1, $1 + x$, $x + x^2$, $x^2 + x^3$.
16. En P_3 escriba el polinomio $4x^2 - x + 5$ en términos de los polinomios base 1, $1 - x$, $(1 - x)^2$, $(1 - x)^3$.
17. En $\mathbb{R}^2$ supongamos que $(\mathbf{x})_{B_1} = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$, donde $B_1 = \left\{\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \end{pmatrix}\right\}$. Escriba $\mathbf{x}$ en términos de la base $B_2 = \left\{\begin{pmatrix} 0 \\ 3 \end{pmatrix}, \begin{pmatrix} 5 \\ -1 \end{pmatrix}\right\}$.
18. En $\mathbb{R}^2$, $(\mathbf{x})_{B_1} = \begin{pmatrix} 4 \\ -1 \end{pmatrix}$, en donde $B_1 = \left\{\begin{pmatrix} 2 \\ -5 \end{pmatrix}, \begin{pmatrix} 7 \\ 3 \end{pmatrix}\right\}$. Escriba $\mathbf{x}$ en términos de $B_2 = \left\{\begin{pmatrix} -2 \\ 1 \end{pmatrix}, \begin{pmatrix} -3 \\ 2 \end{pmatrix}\right\}$.
19. En $\mathbb{R}^3$, $(\mathbf{x})_{B_1} = \begin{pmatrix} 2 \\ -1 \\ 4 \end{pmatrix}$, en donde $B_1 = \left\{\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}\right\}$. Escriba $\mathbf{x}$ en términos de $B_2 = \left\{\begin{pmatrix} 3 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 5 \end{pmatrix}\right\}$.
20. En $P_2 (\mathbf{x})_{B_1} = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}$, en donde $B_1 = \{1 - x, 3x, x^2 - x - 1\}$. Escriba $\mathbf{x}$ en términos de $B_2 = \{3 - 2x, 1 + x, x + x^2\}$.

En los Problemas del 21 al 28 use el Teorema 2 para determinar si el conjunto de vectores dado es linealmente dependiente o independiente.

21. En P_2: $2 + 3x + 5x^2$, $1 - 2x + x^2$, $-1 + 6x^2$
22. En P_2: $-3 + x^2$, $2 - x + 4x^2$, $4 + 2x$
23. En P_2: $x + 4x^2$, $-2 + 2x$, $2 + x + 12x^2$
24. En P_2: $-2 + 4x - 2x^2$, $3 + x$, $6 + 8x$
25. En P_3: $1 + x^2$, $-1 - 3x + 4x^2 + 5x^3$, $2 + 5x - 6x^3$, $4 + 6x + 3x^2 + 7x^3$
26. En M_{22}: $\begin{pmatrix} 2 & 0 \\ 3 & 4 \end{pmatrix}, \begin{pmatrix} -3 & -2 \\ 7 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ -1 & -3 \end{pmatrix}, \begin{pmatrix} 11 & 2 \\ -5 & -5 \end{pmatrix}$
27. En M_{22}: $\begin{pmatrix} 1 & -3 \\ 2 & 4 \end{pmatrix}, \begin{pmatrix} 1 & 4 \\ 5 & 0 \end{pmatrix}, \begin{pmatrix} -1 & 6 \\ -1 & 3 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 3 & 0 \end{pmatrix}$
28. En M_{22}: $\begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} b & c \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} d & e \\ f & 0 \end{pmatrix}, \begin{pmatrix} g & h \\ j & k \end{pmatrix}$, en donde $acfk \neq 0$.



29. En P_n, sean $p_1, p_2, \ldots, p_{n+1}$, $n+1$ polinomios tales que $p_i(0) = 0$ para $i = 1, 2, \ldots, n+1$. Muestre que estos polinomios son linealmente dependientes.

★ □ 30. En el Problema 29 suponga que $p_i^{(j)} = 0$ para $i = 1, 2, \ldots, n+1$ y para alguna j tal que $1 \le j \le n$, donde $p_i^{(j)}$ denota la j-ésima derivada de p_i. Muestre que estos polinomios son linealmente dependientes en P_n.

31. En M_{mn} sean $A_1, A_2, \ldots, A_{mn}$ mn matrices tales que en la posición 1, 1 tienen el número cero. Muestre que estas matrices son linealmente dependientes.

★ 32. Suponga que se giran los ejes x y y un ángulo θ (medido en grados o radianes) en sentido contrario a las manecillas del reloj. Esto nos da nuevos ejes a los que denotaremos con (x', y'). ¿Cuáles son las coordenadas x y y de los vectores **i** y **j** de la base girada?

33. Muestre que la matriz de "cambio de coordenadas" del Problema 32 está dada por

$$A^{-1} = \begin{pmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{pmatrix}.$$

34. Si, en los Problemas 32 y 33, $\theta = \pi/6 = 30°$, escriba el vector $\begin{pmatrix} -4 \\ 3 \end{pmatrix}$ en términos de los nuevos ejes coordenados x' y y'.

35. Si $\theta = \pi/4 = 45°$, escriba el vector $\begin{pmatrix} 2 \\ -7 \end{pmatrix}$ en términos de los nuevos ejes coordenados.

36. Si $\theta = 2\pi/3 = 120°$, escriba $\begin{pmatrix} 4 \\ 5 \end{pmatrix}$ en términos de los nuevos ejes coordenados.

37. Sea $C = (c_{ij})$ una matriz $n \times n$ invertible y sea $B_1 = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ una base para el espacio vectorial V. Sea

$$\mathbf{c}_1 = \begin{pmatrix} c_{11} \\ c_{21} \\ \vdots \\ c_{n1} \end{pmatrix}_{B_1}, \mathbf{c}_2 = \begin{pmatrix} c_{12} \\ c_{22} \\ \vdots \\ c_{n2} \end{pmatrix}_{B_1}, \ldots, \mathbf{c}_n = \begin{pmatrix} c_{1n} \\ c_{2n} \\ \vdots \\ c_{nn} \end{pmatrix}_{B_1}$$

Muestre que $B_2 = \{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n\}$ es una base para V.

38. Sean B_1 y B_2, dos bases para el espacio vectorial V de dimensión n y sea C la matriz de transición de B_1 a B_2. Muestre que C^{-1} es la matriz de transición de B_2 a B_1.

39. Demuestre que $(\mathbf{x})_{B_1} = CA(\mathbf{x})_{B_1}$ para todo $\mathbf{x}$ en un espacio vectorial V si y sólo si $CA = I$. [*Sugerencia:* Sea $\mathbf{x}_i$ el i-ésimo vector de la base B_1. Entonces $(\mathbf{x}_i)_{B_1}$ tiene un 1 en la i-ésima posición y cero en todas las demás. ¿Qué puede decirse de $CA(\mathbf{x}_i)_{B_1}$?]

4.10 Bases ortonormales y proyecciones en $\mathbb{R}^n$

En $\mathbb{R}^n$ ya vimos que n vectores linealmente independientes constituyen una base. La base más usada es la base estándar $E = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$. Estos vectores tienen dos propiedades

i. $\mathbf{e}_i \cdot \mathbf{e}_j = 0$ si $i \ne j$

ii. $\mathbf{e}_i \cdot \mathbf{e}_i = 1$

Definición 1 ***Conjunto ortonormal en $\mathbb{R}^n$.*** El conjunto de vectores $S = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k\}$ en $\mathbb{R}^n$ es un *conjunto ortonormal* si

$$\mathbf{u}_i \cdot \mathbf{u}_j = 0 \quad \text{si } i \ne j \qquad (1)$$

$$\mathbf{u}_i \cdot \mathbf{u}_i = 1 \qquad (2)$$

Si solamente se satisface la Ecuación (1), el conjunto se llama *ortogonal.*

Dado que trabajaremos ampliamente con el producto escalar en esta sección, recordemos algunos resultados básicos (vea el Teorema 1.5.1). Sin hacer mención de ellos de manera explícita, usaremos estos resultados en el resto de esta sección.

Si $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ están en $\mathbb{R}^n$ y si α es un número real, entonces

$$\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u} \qquad (3)$$

$$(\mathbf{u}+\mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{w} \qquad (4)$$

$$\mathbf{u} \cdot (\mathbf{v}+\mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w} \qquad (5)$$

$$(\alpha\mathbf{u}) \cdot \mathbf{v} = \alpha(\mathbf{u} \cdot \mathbf{v}) \qquad (6)$$

$$\mathbf{u} \cdot (\alpha\mathbf{v}) = \alpha(\mathbf{u} \cdot \mathbf{v}) \qquad (7)$$

Definición 2 ***Norma o longitud de un vector.*** Veamos otra definición útil.

Si $\mathbf{v} \in \mathbb{R}^n$ entonces la *norma* o *longitud* de $\mathbf{v}$, denotada como $|\mathbf{v}|$ está dada por

$$|\mathbf{v}| = \sqrt{\mathbf{v} \cdot \mathbf{v}} \qquad (8)$$

Nota. Si $\mathbf{v} = (x_1, x_2, \ldots, x_n)$, entonces $\mathbf{v} \cdot \mathbf{v} = x_1^2 + x_2^2 + \cdots + x_n^2$. Esto significa que

$$\mathbf{v} \cdot \mathbf{v} \ge 0 \quad \text{y} \quad \mathbf{v} \cdot \mathbf{v} = 0 \quad \text{si y sólo si } \mathbf{v} = \mathbf{0} \qquad (9)$$

De tal forma que podemos extraer raíz cuadrada en (8) y obtenemos

$$|\mathbf{v}| = \sqrt{\mathbf{v} \cdot \mathbf{v}} \ge 0 \quad \text{para todo } \mathbf{v} \in \mathbb{R}^n \qquad (10)$$

$$|\mathbf{v}| = 0 \quad \text{si y sólo si } \mathbf{v} = \mathbf{0} \qquad (11)$$

Ejemplo 1 Sea $\mathbf{v} = (x, y) \in \mathbb{R}^2$. Entonces $|\mathbf{v}| = \sqrt{x^2+y^2}$ es simplemente la definición usual de longitud de un vector en el plano (vea la Ecuación 3.1.1).

Ejemplo 2 Si $\mathbf{v} = (x, y, z) \in \mathbb{R}^3$ entonces $|\mathbf{v}| = \sqrt{x^2+y^2+z^2}$ como en la Sección 3.3.

Ejemplo 3 Si $\mathbf{v}=(2,-1,3,4,-6)\in\mathbb{R}^5$, entonces $|\mathbf{v}|=\sqrt{4+1+9+16+36}=\sqrt{66}$.

Ahora podemos enunciar la Definición 1 de la siguiente manera:

> Un conjunto de vectores es ortonormal si cualquier par de esos vectores es ortogonal y cada uno tiene longitud 1.

Es relativamente fácil trabajar con conjuntos ortonormales. En el Capítulo 5 veremos un ejemplo de esto. Ahora demostraremos que cualquier conjunto ortogonal finito de vectores distintos de cero es linealmente independiente.

Teorema 1 Si $S=\{\mathbf{v}_1,\mathbf{v}_2,\ldots,\mathbf{v}_k\}$ es un conjunto ortogonal de vectores no nulos, entonces S es linealmente independiente.

Demostración Supongamos que $c_1\mathbf{v}_1+c_2\mathbf{v}_2+\cdots+c_n\mathbf{v}_n=\mathbf{0}$. Entonces para $i=1,2,\ldots,k$,

$$\begin{aligned}0=\mathbf{0}\cdot\mathbf{v}_i&=(c_1\mathbf{v}_1+c_2\mathbf{v}_2+\cdots+c_i\mathbf{v}_i+\cdots+c_n\mathbf{v}_n)\cdot\mathbf{v}_i\\&=c_1(\mathbf{v}_1\cdot\mathbf{v}_i)+c_2(\mathbf{v}_2\cdot\mathbf{v}_i)+\cdots+c_i(\mathbf{v}_i\cdot\mathbf{v}_i)+\cdots+c_n(\mathbf{v}_n\cdot\mathbf{v}_i)\\&=c_10+c_20+\cdots+c_i|\mathbf{v}_i|^2+\cdots+c_n0=c_i|\mathbf{v}_i|^2\end{aligned}$$

Puesto que $\mathbf{v}_i\neq 0$ por hipótesis, $|\mathbf{v}_i|^2>0$ y tenemos que $c_i=0$. Como esto es cierto para $i=1,2,\ldots,k$, la demostración está completa. ■

Veamos ahora que *cualquier* base en $\mathbb{R}^n$ puede ser "convertida" a una base ortonormal. El método que se describe a continuación se conoce como el *proceso de ortonormalización de Gram-Schmidt.**

Teorema 2 **Proceso de ortonormalización de Gram-Schmidt.** Sea H un subespacio de dimensión m de $\mathbb{R}^n$. Entonces H tiene una base ortonormal.†

Demostración Sea $S=\{\mathbf{v}_1,\mathbf{v}_2,\ldots,\mathbf{v}_m\}$ una base para H. Demostraremos el teorema construyendo una base ortonormal a partir de los vectores de S. Antes de dar los pasos para esta construcción, notemos que un conjunto de vectores linealmente independiente *no* contiene al vector cero (vea el Problema 21).

Paso 1. Sea

$$\mathbf{u}_1=\frac{\mathbf{v}_1}{|\mathbf{v}_1|}\tag{12}$$

* Jörgen Pederson Gram (1850-1916) fue un actuario danés que estuvo muy interesado en la ciencia metrológica. Erhardt Schmidt (1876-1959) fue un matemático alemán.

† Notemos que H puede ser $\mathbb{R}^n$ en este teorema. Esto es, $\mathbb{R}^n$ tiene una base ortonormal.

Entonces $\mathbf{u}_1\cdot\mathbf{u}_1=(\mathbf{v}_1/|\mathbf{v}_1|)\cdot(\mathbf{v}_1/|\mathbf{v}_1|)=(1/|\mathbf{v}_1|^2)(\mathbf{v}_1\cdot\mathbf{v}_1)=1$, de donde $|\mathbf{u}_1|=1$.

Paso 2. Sea

$$\mathbf{v}_2'=\mathbf{v}_2-(\mathbf{v}_2\cdot\mathbf{u}_1)\mathbf{u}_1\tag{13}$$

Entonces $\mathbf{v}_2'\cdot\mathbf{u}_1=\mathbf{v}_2\cdot\mathbf{u}_1-(\mathbf{v}_2\cdot\mathbf{u}_1)(\mathbf{u}_1\cdot\mathbf{u}_1)=\mathbf{v}_2\cdot\mathbf{u}_1-\mathbf{v}_2\cdot\mathbf{u}_1=0$, de tal forma que $\mathbf{v}_2'$ es ortogonal a $\mathbf{u}_1$. Además por el Teorema 1, $\mathbf{u}_1$ y $\mathbf{v}_2'$ son linealmente independientes, y por el Problema 21, $\mathbf{v}_2'\neq\mathbf{0}$.

Paso 3. Sea

$$\mathbf{u}_2=\frac{\mathbf{v}_2'}{|\mathbf{v}_2'|}\tag{14}$$

claramente $\{\mathbf{u}_1,\mathbf{u}_2\}$ es un conjunto ortonormal.

Supongamos ahora que los vectores $\mathbf{u}_1,\mathbf{u}_2,\ldots,\mathbf{u}_k$ $(k<m)$ ya han sido construidos y que forman un conjunto ortonormal. Mostraremos cómo construir $\mathbf{u}_{k+1}$.

Paso 4. Sea

$$\mathbf{v}_{k+1}'=\mathbf{v}_{k+1}-(\mathbf{v}_{k+1}\cdot\mathbf{u}_1)\mathbf{u}_1-(\mathbf{v}_{k+1}\cdot\mathbf{u}_2)\mathbf{u}_2-\cdots-(\mathbf{v}_{k+1}\cdot\mathbf{u}_k)\mathbf{u}_k\tag{15}$$

Entonces, para $i=1,2,\ldots,k$,

$$\begin{aligned}\mathbf{v}_{k+1}'\cdot\mathbf{u}_i=\mathbf{v}_{k+1}\cdot\mathbf{u}_i&-(\mathbf{v}_{k+1}\cdot\mathbf{u}_1)(\mathbf{u}_1\cdot\mathbf{u}_i)-(\mathbf{v}_{k+1}\cdot\mathbf{u}_2)(\mathbf{u}_2\cdot\mathbf{u}_i)\\&-\cdots-(\mathbf{v}_{k+1}\cdot\mathbf{u}_i)(\mathbf{u}_i\cdot\mathbf{u}_i)-\cdots-(\mathbf{v}_{k+1}\cdot\mathbf{u}_k)(\mathbf{u}_k\cdot\mathbf{u}_i)\end{aligned}$$

Pero $\mathbf{u}_j\cdot\mathbf{u}_i=0$ si $j\neq i$ y $\mathbf{u}_i\ \mathbf{u}_i=1$. Así

$$\mathbf{v}_{k+1}'\cdot\mathbf{u}_i=\mathbf{v}_{k+1}\cdot\mathbf{u}_i-\mathbf{v}_{k+1}\cdot\mathbf{u}_i=0$$

Por lo tanto $\{\mathbf{u}_1,\mathbf{u}_2,\ldots,\mathbf{u}_k,\mathbf{v}_{k+1}'\}$ es un conjunto ortogonal y linealmente independiente. Además $\mathbf{v}_{k+1}'\neq\mathbf{0}$.

Paso 5. Sea $\mathbf{u}_{k+1}=\mathbf{v}_{k+1}'/|\mathbf{v}_{k+1}'|$. Claramente $\{\mathbf{u}_1,\mathbf{u}_2,\ldots,\mathbf{u}_k,\mathbf{u}_{k+1}\}$ es un conjunto ortonormal y así continuamos este proceso hasta que $k+1=m$; con esto completamos la demostración. ■

Ejemplo 4 Encontrar una base ortonormal para el conjunto de vectores en $\mathbb{R}^3$ que están sobre el plano $\pi=\left\{\begin{pmatrix}x\\y\\z\end{pmatrix}:\ 2x-y+3z=0\right\}$.

Solución Tal y como vimos en el Ejemplo 4.6.3 una base para este subespacio de dimensión dos es $\mathbf{v}_1=\begin{pmatrix}1\\2\\0\end{pmatrix}$ y $\mathbf{v}_2=\begin{pmatrix}0\\3\\1\end{pmatrix}$. Entonces $|\mathbf{v}_1|=\sqrt{5}$ y $\mathbf{u}_1=\mathbf{v}_1/|\mathbf{v}_1|=$

$\begin{pmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \\ 0 \end{pmatrix}$. Para continuar, definimos

$$\mathbf{v}_2' = \mathbf{v}_2 - (\mathbf{v}_2 \cdot \mathbf{u}_1)\mathbf{u}_1$$
$$= \begin{pmatrix} 0 \\ 3 \\ 1 \end{pmatrix} - (6/\sqrt{5})\begin{pmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 3 \\ 1 \end{pmatrix} - \begin{pmatrix} 6/5 \\ 12/5 \\ 0 \end{pmatrix} = \begin{pmatrix} -6/5 \\ 3/5 \\ 1 \end{pmatrix}$$

Finalmente tenemos $|\mathbf{v}_2'| = \sqrt{70/25} = \sqrt{70}/5$, de tal forma que $\mathbf{u}_2 = \mathbf{v}_2'/|\mathbf{v}_2'| = (5/\sqrt{70})\begin{pmatrix} -6/5 \\ 3/5 \\ 1 \end{pmatrix} = \begin{pmatrix} -6/\sqrt{70} \\ 3/\sqrt{70} \\ 5/\sqrt{70} \end{pmatrix}$. Así la base ortonormal es $\left\{ \begin{pmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \\ 0 \end{pmatrix}, \begin{pmatrix} -6/\sqrt{70} \\ 3/\sqrt{70} \\ 5/\sqrt{70} \end{pmatrix} \right\}$.

Para verificar esta solución, notemos (i) los vectores son ortogonales, (ii) cada uno tiene longitud 1 y (iii) cada uno satisface a $2x - y + 3z = 0$.

Observación. Podemos ver geométricamente lo que estamos haciendo aquí. Primero notemos que

$$\mathbf{v}_2' = \mathbf{v}_2 - (\mathbf{v}_2 \cdot \mathbf{u}_1)\mathbf{u}_1 = \mathbf{v}_2 - \left(\mathbf{v}_2 \cdot \frac{\mathbf{v}_1}{|\mathbf{v}_1|}\right)\left(\frac{\mathbf{v}_1}{|\mathbf{v}_1|}\right) = \mathbf{v}_2 - \frac{(\mathbf{v}_2 \cdot \mathbf{v}_1)}{|\mathbf{v}_1|^2}\mathbf{v}_1$$

Pero, de las Definiciones 3.2.4 (en $\mathbb{R}^2$) y 3.3.3 (en $\mathbb{R}^3$), $[(\mathbf{v}_2 \cdot \mathbf{v}_1)/|\mathbf{v}_1|^2]\mathbf{v}_1$ es la proyección de $\mathbf{v}_2$ sobre $\mathbf{v}_1$. Además de la Figura 3.13, el vector $\mathbf{v}_2' = \mathbf{v}_2 - \text{proy}_{\mathbf{v}_1}\mathbf{v}_2$ es un vector ortogonal a $\mathbf{v}_1$.

De esta forma podemos ver que el proceso que hemos descrito aquí es, en cierto sentido, una generalización de la noción de proyección en $\mathbb{R}^2$ y $\mathbb{R}^3$.

Ejemplo 5 Construya una base ortonormal en $\mathbb{R}^3$ a partir de la base $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\} = \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\}$.

Solución Tenemos que $|\mathbf{v}_1| = \sqrt{2}$, de tal forma que $\mathbf{u}_1 = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix}$. Entonces

$$\mathbf{v}_2' = \mathbf{v}_2 - (\mathbf{v}_2 \cdot \mathbf{u}_1)\mathbf{u}_1 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} - \frac{1}{\sqrt{2}}\begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} - \begin{pmatrix} 1/2 \\ 1/2 \\ 0 \end{pmatrix} = \begin{pmatrix} -1/2 \\ 1/2 \\ 1 \end{pmatrix}$$

Como $|\mathbf{v}_2'| = \sqrt{3/2}$, $|\mathbf{u}_2| = \sqrt{2/3}\begin{pmatrix} -1/2 \\ 1/2 \\ 1 \end{pmatrix} = \begin{pmatrix} -1/\sqrt{6} \\ 1/\sqrt{6} \\ 2/\sqrt{6} \end{pmatrix}$. Continuando, tenemos

$$\mathbf{v}_3' = \mathbf{v}_3 - (\mathbf{v}_3 \cdot \mathbf{u}_1)\mathbf{u}_1 - (\mathbf{v}_3 \cdot \mathbf{u}_2)\mathbf{u}_2$$
$$= \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} - \frac{1}{\sqrt{2}}\begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix} - \frac{1}{\sqrt{6}}\begin{pmatrix} -1/\sqrt{6} \\ 1/\sqrt{6} \\ 2/\sqrt{6} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} - \begin{pmatrix} 1/2 \\ 1/2 \\ 0 \end{pmatrix} - \begin{pmatrix} -1/6 \\ 1/6 \\ 2/6 \end{pmatrix} = \begin{pmatrix} 2/3 \\ -2/3 \\ 2/3 \end{pmatrix}$$

Finalmente $|\mathbf{v}_3'| = \sqrt{\frac{12}{9}} = 2/\sqrt{3}$, de tal forma que $\mathbf{u}_3 = \frac{\sqrt{3}}{2}\begin{pmatrix} 2/3 \\ -2/3 \\ 2/3 \end{pmatrix} = \begin{pmatrix} 1/\sqrt{3} \\ -1/\sqrt{3} \\ 1/\sqrt{3} \end{pmatrix}$.

Así, la base ortonormal es $\left\{ \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix}, \begin{pmatrix} -1/\sqrt{6} \\ 1/\sqrt{6} \\ 2/\sqrt{6} \end{pmatrix}, \begin{pmatrix} 1/\sqrt{3} \\ -1/\sqrt{3} \\ 1/\sqrt{3} \end{pmatrix} \right\}$. Como en el ejemplo anterior, este resultado debe comprobarse.

Definimos ahora una clase nueva de matrices que será muy útil en los últimos capítulos.

Definición 3 ***Matriz ortogonal.*** Se dice que una matriz Q de tamaño $n \times n$ es *ortogonal* si Q es invertible y

$$Q^{-1} = Q^t \qquad (16)$$

Según el siguiente teorema no es muy difícil encontrar matrices ortogonales.

Teorema 3 Una matriz Q de tamaño $n \times n$ es ortogonal si y sólo si las columnas de Q forman una base ortonormal para $\mathbb{R}^n$.

Demostración Sea

$$Q = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

Entonces

$$Q^t = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{pmatrix}$$

Sea $B = (b_{ij}) = Q^tQ$. Entonces

$$b_{ij} = a_{1i}a_{1j} + a_{2i}a_{2j} + \cdots + a_{ni}a_{nj} = \mathbf{c}_i \cdot \mathbf{c}_j \qquad (17)$$

donde $\mathbf{c}_i$ denota la i-ésima columna de Q. Si las columnas de Q son ortonormales, entonces

$$b_{ij} = \begin{cases} 0 & \text{si} \quad i \neq j \\ 1 & \text{si} \quad i = j \end{cases} \tag{18}$$

Es decir, $B = I$. Recíprocamente, si $Q^t = Q^{-1}$, entonces $B = I$, de donde (18) se satisface y (17) muestra que las columnas de Q son ortonormales. Esto completa la demostración. ■

Ejemplo 6 Del Ejemplo 5, los vectores $\begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix}, \begin{pmatrix} -1/\sqrt{6} \\ 1/\sqrt{6} \\ 2/\sqrt{6} \end{pmatrix}, \begin{pmatrix} 1/\sqrt{3} \\ -1/\sqrt{3} \\ 1/\sqrt{3} \end{pmatrix}$ forman una base ortonormal en $\mathbb{R}^3$. Así, la matriz $Q = \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 1/\sqrt{2} & 1/\sqrt{6} & -1/\sqrt{3} \\ 0 & 2/\sqrt{6} & 1/\sqrt{3} \end{pmatrix}$ es una matriz ortogonal. Para verificar esto, notemos que

$$Q^tQ = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ -1/\sqrt{6} & 1/\sqrt{6} & 2/\sqrt{6} \\ 1/\sqrt{3} & -1/\sqrt{3} & 1/\sqrt{3} \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 1/\sqrt{2} & 1/\sqrt{6} & -1/\sqrt{3} \\ 0 & 2/\sqrt{6} & 1/\sqrt{3} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

En la demostración del Teorema 2 definimos $\mathbf{v}_2' = \mathbf{v}_2 - (\mathbf{v}_2 \cdot \mathbf{u}_1)\mathbf{u}_1$. Pero, como ya hemos visto, $(\mathbf{v}_2 \cdot \mathbf{u}_1)\mathbf{u}_1 = \text{proy}_{\mathbf{u}_1}\mathbf{v}_2$ (ya que $|\mathbf{u}_1|^2 = 1$). Ahora extendamos el concepto de proyección sobre un vector a la de proyección sobre un subespacio.

Definición 4 ***Proyección ortogonal.*** Sea H un subespacio de $\mathbb{R}^n$ con base ortonormal $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k\}$. Si $\mathbf{v} \in \mathbb{R}^n$, entonces la *proyección ortogonal* de $\mathbf{v}$ sobre H, denotada por $\text{proy}_H\mathbf{v}$, está dada por

$$\text{proy}_H\,\mathbf{v} = (\mathbf{v} \cdot \mathbf{u}_1)\mathbf{u}_1 + (\mathbf{v} \cdot \mathbf{u}_2)\mathbf{u}_2 + \cdots + (\mathbf{v} \cdot \mathbf{u}_k)\mathbf{u}_k \tag{19}$$

Notemos que $\text{proy}_H\,\mathbf{v} \in H$.

Antes de dar ejemplos debemos saber que está bien definida una proyección ortogonal. Con esto queremos decir que la definición de $\text{proy}_H\mathbf{v}$ es independiente de la base ortonormal elegida en H. El siguiente teorema toma en cuenta este problema. Su demostración se difiere hacia el final de esta sección.

Teorema 4 Sea H un subespacio de $\mathbb{R}^n$. Supóngase que H tiene dos bases ortonormales, $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k\}$ y $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_k\}$. Sea $\mathbf{v}$ un vector en $\mathbb{R}^n$. Entonces

$$(\mathbf{v}\cdot\mathbf{u}_1)\mathbf{u}_1 + (\mathbf{v}\cdot\mathbf{u}_2)\mathbf{u}_2 + \cdots + (\mathbf{v}\cdot\mathbf{u}_k)\mathbf{u}_k = (\mathbf{v}\cdot\mathbf{w}_1)\mathbf{w}_1 + (\mathbf{v}\cdot\mathbf{w}_2)\mathbf{w}_2 + \cdots + (\mathbf{v}\cdot\mathbf{w}_k)\mathbf{w}_k \tag{20}$$

Ejemplo 7 Encuentre $\text{proy}_\pi\mathbf{v}$, donde π es el plano $\left\{\begin{pmatrix} x \\ y \\ z \end{pmatrix}: 2x - y + 3z = 0\right\}$ y asimismo $\mathbf{v}$ es el vector $\begin{pmatrix} 3 \\ -2 \\ 4 \end{pmatrix}$.

Solución Del Ejemplo 4, una base ortonormal para π es $\mathbf{u}_1 = \begin{pmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \\ 0 \end{pmatrix}$ y $\mathbf{u}_2 = \begin{pmatrix} -6/\sqrt{70} \\ 3/\sqrt{70} \\ 5/\sqrt{70} \end{pmatrix}$. Entonces

$$\text{proy}_\pi\,\mathbf{v} = \left[\begin{pmatrix} 3 \\ -2 \\ 4 \end{pmatrix} \cdot \begin{pmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \\ 0 \end{pmatrix}\right]\begin{pmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \\ 0 \end{pmatrix} + \left[\begin{pmatrix} 3 \\ -2 \\ 4 \end{pmatrix} \cdot \begin{pmatrix} -6/\sqrt{70} \\ 3/\sqrt{70} \\ 5/\sqrt{70} \end{pmatrix}\right]\begin{pmatrix} -6/\sqrt{70} \\ 3/\sqrt{70} \\ 5/\sqrt{70} \end{pmatrix}$$

$$= -\frac{1}{\sqrt{5}}\begin{pmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \\ 0 \end{pmatrix} - \frac{4}{\sqrt{70}}\begin{pmatrix} -6/\sqrt{70} \\ 3/\sqrt{70} \\ 5/\sqrt{70} \end{pmatrix} = \begin{pmatrix} -1/5 \\ -2/5 \\ 0 \end{pmatrix} + \begin{pmatrix} 24/70 \\ -12/70 \\ -20/70 \end{pmatrix} = \begin{pmatrix} 1/7 \\ -4/7 \\ -2/7 \end{pmatrix}$$

El concepto de proyección nos da una forma sencilla de escribir un vector en $\mathbb{R}^n$ en términos de una base ortonormal.

Teorema 5 Sea $B = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ una base ortonormal de $\mathbb{R}^n$ y sea $\mathbf{v} \in \mathbb{R}^n$. Entonces

$$\mathbf{v} = (\mathbf{v} \cdot \mathbf{u}_1)\mathbf{u}_1 + (\mathbf{v} \cdot \mathbf{u}_2)\mathbf{u}_2 + \cdots + (\mathbf{v} \cdot \mathbf{u}_n)\mathbf{u}_n \tag{21}$$

Es decir, $\mathbf{v} = \text{proy}_{\mathbb{R}^n}\,\mathbf{v}$.

Demostración Dado que B es una base, podemos escribir $\mathbf{v}$ de manera única como $\mathbf{v} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_n\mathbf{u}_n$. Entonces

$$\mathbf{v} \cdot \mathbf{u}_i = c_1(\mathbf{u}_1 \cdot \mathbf{u}_i) + c_2(\mathbf{u}_2 \cdot \mathbf{u}_i) + \cdots + c_i(\mathbf{u}_i \cdot \mathbf{u}_i) + \cdots + c_n(\mathbf{u}_n \cdot \mathbf{u}_i) = c_i$$

puesto que los $\mathbf{u}_i$ son ortonormales. Como esto es válido para $i = 1, 2, \ldots, n$, hemos terminado la demostración. ■

Ejemplo 8 Escriba el vector $\begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix}$ en $\mathbb{R}^3$ en términos de la base ortonormal que sigue:

$$\left\{\begin{pmatrix}1/\sqrt{2}\\1/\sqrt{2}\\0\end{pmatrix},\begin{pmatrix}-1/\sqrt{6}\\1/\sqrt{6}\\2/\sqrt{6}\end{pmatrix},\begin{pmatrix}1/\sqrt{3}\\-1/\sqrt{3}\\1/\sqrt{3}\end{pmatrix}\right\}.$$

Solución

$$\begin{pmatrix}2\\-1\\3\end{pmatrix}=\left[\begin{pmatrix}2\\-1\\3\end{pmatrix}\cdot\begin{pmatrix}1/\sqrt{2}\\1/\sqrt{2}\\0\end{pmatrix}\right]\begin{pmatrix}1/\sqrt{2}\\1/\sqrt{2}\\0\end{pmatrix}+\left[\begin{pmatrix}2\\-1\\3\end{pmatrix}\cdot\begin{pmatrix}-1/\sqrt{6}\\1/\sqrt{6}\\2/\sqrt{6}\end{pmatrix}\right]\begin{pmatrix}-1/\sqrt{6}\\1/\sqrt{6}\\2/\sqrt{6}\end{pmatrix}$$

$$+\left[\begin{pmatrix}2\\-1\\3\end{pmatrix}\cdot\begin{pmatrix}1/\sqrt{3}\\-1/\sqrt{3}\\1/\sqrt{3}\end{pmatrix}\right]\begin{pmatrix}1/\sqrt{3}\\-1/\sqrt{3}\\1/\sqrt{3}\end{pmatrix}=\frac{1}{\sqrt{2}}\begin{pmatrix}1/\sqrt{2}\\1/\sqrt{2}\\0\end{pmatrix}+\frac{3}{\sqrt{6}}\begin{pmatrix}-1/\sqrt{6}\\1/\sqrt{6}\\2/\sqrt{6}\end{pmatrix}+\frac{6}{\sqrt{3}}\begin{pmatrix}1/\sqrt{3}\\-1/\sqrt{3}\\1/\sqrt{3}\end{pmatrix}$$

Podemos usar el Teorema 5 para demostrar el Teorema 4.

Demostración del Teorema 4 Elíjanse vectores $\mathbf{u}_{k+1}, \mathbf{u}_{k+2}, \ldots, \mathbf{u}_n$ tales que $B_1 = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k, \mathbf{u}_{k+1}, \ldots, \mathbf{u}_n\}$ sea una base ortonormal de $\mathbb{R}^n$ (esto siempre se puede efectuar como se hizo en el Teorema 2). Entonces $B_2 = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_k, \mathbf{u}_{k+1}, \mathbf{u}_{k+2}, \ldots, \mathbf{u}_n\}$ también es una base ortonormal de $\mathbb{R}^n$. Para ver esto, primero nótese que ninguno de los vectores $\mathbf{u}_{k+1}, \mathbf{u}_{k+2}, \ldots, \mathbf{u}_n$ puede ser escrito como combinación lineal de $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_k$ porque ninguno de estos vectores está en H y$\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_k\}$ es una base de H. Por lo tanto, B_2 es una base de $\mathbb{R}^n$ porque contiene n vectores linealmente independientes. La ortonormalidad de los vectores en B_2 se sigue de la manera en que fueron elegidos ($\mathbf{u}_{k+j}$ es ortogonal a todo vector en H para toda $j = 1, 2, \ldots, n-k$). Sea $\mathbf{v}$ un vector en $\mathbb{R}^n$. Entonces por el Teorema 5 (Ecuación 21)),

$$\begin{aligned}\mathbf{v} &= (\mathbf{v}\cdot\mathbf{u}_1)\mathbf{u}_1 + (\mathbf{v}\cdot\mathbf{u}_2)\mathbf{u}_2 + \cdots + (\mathbf{v}\cdot\mathbf{u}_k)\mathbf{u}_k + (\mathbf{v}\cdot\mathbf{u}_{k+1})\mathbf{u}_{k+1} + \cdots + (\mathbf{v}\cdot\mathbf{u}_n)\mathbf{u}_n\\ &= (\mathbf{v}\cdot\mathbf{w}_1)\mathbf{w}_1 + (\mathbf{v}\cdot\mathbf{w}_2)\mathbf{w}_2 + \cdots + (\mathbf{v}\cdot\mathbf{w}_k)\mathbf{w}_k + (\mathbf{v}\cdot\mathbf{u}_{k+1})\mathbf{u}_{k+1} + \cdots + (\mathbf{v}\cdot\mathbf{u}_n)\mathbf{u}_n\end{aligned} \tag{22}$$

La Ecuación (20) se deduce de la Ecuación (22). ■

Definición 5 ***Complemento ortogonal.*** Sea H un subespacio de $\mathbb{R}^n$. Entonces el *complemento ortogonal* de H, denotado por $H^\perp$, está dado por

$$H^\perp = \{\mathbf{x}\in\mathbb{R}^n: \quad \mathbf{x}\cdot\mathbf{h}=0 \quad \text{para todo} \quad \mathbf{h}\in H\}$$

Teorema 6 Si H es un subespacio de $\mathbb{R}^n$, entonces:

i. $H^\perp$ es un subespacio de $\mathbb{R}^n$.
ii. $H \cap H^\perp = \{\mathbf{0}\}$.
iii. $\dim H^\perp = n - \dim H$.

Demostración

i. Si $\mathbf{x}$ y $\mathbf{y}$ están en $H^\perp$ y si $\mathbf{h} \in H$, entonces $(\mathbf{x}+\mathbf{y})\cdot\mathbf{h} = \mathbf{x}\cdot\mathbf{h} + \mathbf{y}\cdot\mathbf{h} = 0 + 0 = 0$ y $(\alpha\mathbf{x}\cdot\mathbf{h}) = \alpha(\mathbf{x}\cdot\mathbf{h}) = 0$, por lo que $H^\perp$ es un subespacio.

ii. Si $\mathbf{x} \in H \cap H^\perp$, entonces $\mathbf{x}\cdot\mathbf{x} = 0$, por lo que $\mathbf{x} = \mathbf{0}$, lo cual muestra que $H \cap H^\perp = \{\mathbf{0}\}$.

iii. Sea $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k\}$ una base ortonormal de H. Por el resultado del Problema 4.6.27 esta base puede extenderse a una base B para $\mathbb{R}^n$: $B = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k, \mathbf{v}_{k+1}, \ldots, \mathbf{v}_n\}$. Por medio del proceso de Gram-Schmidt, podemos transformar B en una base ortonormal para $\mathbb{R}^n$. Tal y como vimos en la demostración del Teorema 2, la base $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k$ ya ortonormal, no cambia y de esta manera obtenemos la base ortonormal $B_1 = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k, \mathbf{u}_{k+1}, \ldots, \mathbf{u}_n\}$. Para completar la demostración sólo nos falta mostrar que $\{\mathbf{u}_{k+1}, \ldots, \mathbf{u}_n\}$ es una base para $H^\perp$ y dado que los $\mathbf{u}_i$ son independientes, debemos mostrar que generan $H^\perp$. Sea $\mathbf{x} \in H^\perp$; entonces, por el Teorema 5,

$$\begin{aligned}\mathbf{x} = (\mathbf{x}\cdot\mathbf{u}_1)\mathbf{u}_1 + (\mathbf{x}\cdot\mathbf{u}_2)\mathbf{u}_2 + \cdots + (\mathbf{x}\cdot\mathbf{u}_k)\mathbf{u}_k\\ + (\mathbf{x}\cdot\mathbf{u}_{k+1})\mathbf{u}_{k+1} + \cdots + (\mathbf{x}\cdot\mathbf{u}_n)\mathbf{u}_n\end{aligned}$$

Pero $(\mathbf{x}\cdot\mathbf{u}_i) = 0$ para $i = 1, 2, \ldots, k$, puesto que $\mathbf{x} \in H^\perp$ y $\mathbf{u}_i \in H$. Por lo tanto, $\mathbf{x} = (\mathbf{x}\cdot\mathbf{u}_{k+1})\mathbf{u}_{k+1} + \cdots + (\mathbf{x}\cdot\mathbf{u}_n)\mathbf{u}_n$. Esto muestra que $\{\mathbf{u}_{k+1}, \ldots, \mathbf{u}_n\}$ es una base para $H^\perp$, lo cual implica que $\dim H^\perp = n - k$. ■

Los espacios H y $H^\perp$ nos permiten "descomponer" cualquier vector en $\mathbb{R}^n$.

Teorema 7 ***Teorema de la proyección.*** Sea H un subespacio de $\mathbb{R}^n$ y sea $\mathbf{v} \in \mathbb{R}^n$. Entonces existe un único par de vectores $\mathbf{h}$ y $\mathbf{p}$ tal que $\mathbf{h} \in H$, $\mathbf{p} \in H^\perp$ y

$$\mathbf{v} = \mathbf{h} + \mathbf{p} = \text{proy}_H\mathbf{v} + \text{proy}_{H^\perp}\mathbf{v} \tag{23}$$

Demostración Sean $\mathbf{h} = \text{proy}_H\mathbf{v}$ y $\mathbf{p} = \mathbf{v} - \mathbf{h}$. Por la Definición 4, se tiene que $\mathbf{h} \in H$. Mostraremos ahora que $\mathbf{p} \in H^\perp$. Sea $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k\}$ una base para H. Entonces

$$\mathbf{h} = (\mathbf{v}\cdot\mathbf{u}_1)\mathbf{u}_1 + (\mathbf{v}\cdot\mathbf{u}_2)\mathbf{u}_2 + \cdots + (\mathbf{v}\cdot\mathbf{u}_k)\mathbf{u}_k$$

Sea $\mathbf{x}$ un vector en H. Existen constantes $\alpha_1, \alpha_2, \ldots, \alpha_k$ tales que

$$\mathbf{x} = \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_k\mathbf{u}_k$$

Entonces

$$\begin{aligned}\mathbf{p}\cdot\mathbf{x} = (\mathbf{v}-\mathbf{h})\cdot\mathbf{x} = [\mathbf{v} - (\mathbf{v}\cdot\mathbf{u}_1)\mathbf{u}_1 - (\mathbf{v}\cdot\mathbf{u}_2)\mathbf{u}_2 - \cdots - (\mathbf{v}\cdot\mathbf{u}_k)\mathbf{u}_k]\\ \cdot[\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_k\mathbf{u}_k]\end{aligned} \tag{24}$$

Como $\mathbf{u}_i \cdot \mathbf{u}_j = \begin{cases}0, & i \neq j\\ 1, & i = j\end{cases}$, es fácil verificar que el producto escalar en (24) está dado por

$$\mathbf{p}\cdot\mathbf{x} = \sum_{i=1}^{k}\alpha_i(\mathbf{v}\cdot\mathbf{u}_i) - \sum_{i=1}^{k}\alpha_i(\mathbf{v}\cdot\mathbf{u}_i) = 0$$

Así pues $\mathbf{p}\cdot\mathbf{x} = 0$ para todo $\mathbf{x} \in H$, lo que significa que $\mathbf{p} \in H^\perp$. Para demostrar que $\mathbf{p} = \text{proy}_{H^\perp}\mathbf{v}$ extendemos $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k\}$ a una base ortonormal

de $\mathbb{R}^n$: $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k, \mathbf{u}_{k+1}, \ldots, \mathbf{u}_n\}$. Entonces $\{\mathbf{u}_{k+1}, \ldots, \mathbf{u}_n\}$ es una base de $H^\perp$ y, por el Teorema 5,

$$\begin{aligned}\mathbf{v} &= (\mathbf{v}\cdot\mathbf{u}_1)\mathbf{u}_1 + (\mathbf{v}\cdot\mathbf{u}_2)\mathbf{u}_2 + \cdots + (\mathbf{v}\cdot\mathbf{u}_k)\mathbf{u}_k + (\mathbf{v}\cdot\mathbf{u}_{k+1})\mathbf{u}_{k+1} \\ &\quad + \cdots + (\mathbf{v}\cdot\mathbf{u}_n)\mathbf{u}_n \\ &= \text{proy}_H\,\mathbf{v} + \text{proy}_{H^\perp}\mathbf{v} \quad \text{(por la Definición 4)}\end{aligned}$$

Esto demuestra la Ecuación (23). Para demostrar la unicidad, supongamos que $\mathbf{v}=\mathbf{h}_1-\mathbf{p}_1=\mathbf{h}_2-\mathbf{p}_2$, donde $\mathbf{h}_1, \mathbf{h}_2 \in H$ y $\mathbf{p}_1, \mathbf{p}_2 \in H^\perp$. Entonces $\mathbf{h}_1-\mathbf{h}_2=\mathbf{p}_1-\mathbf{p}_2$. Pero $\mathbf{h}_1-\mathbf{h}_2 \in H$ y $\mathbf{p}_1-\mathbf{p}_2 \in H^\perp$, por lo que $\mathbf{h}_1-\mathbf{h}_2 \in H \cap H^\perp = \{\mathbf{0}\}$. Por lo tanto, $\mathbf{h}_1-\mathbf{h}_2=\mathbf{0}$ y $\mathbf{p}_1-\mathbf{p}_2=\mathbf{0}$, lo que completa la demostración. ■

Ejemplo 9 En $\mathbb{R}^3$ sea $\pi = \left\{\begin{pmatrix}x\\y\\z\end{pmatrix}: \; 2x-y+3z=0\right\}$. Escriba el vector $\begin{pmatrix}3\\-2\\4\end{pmatrix}$ como $\mathbf{h}+\mathbf{p}$ donde $\mathbf{h}\in H$ y $\mathbf{p}\in H^\perp$.

Solución Una base ortonormal para π es $B_1 = \left\{\begin{pmatrix}1/\sqrt{5}\\2/\sqrt{5}\\0\end{pmatrix}, \begin{pmatrix}-6/\sqrt{70}\\3/\sqrt{70}\\5/\sqrt{70}\end{pmatrix}\right\}$ y, del Ejemplo 7, $\text{proy}_\pi\,\mathbf{v} = \begin{pmatrix}1/7\\-4/7\\-2/7\end{pmatrix} \in \pi$. Entonces

$$\mathbf{p}=\mathbf{v}-\mathbf{h}=\begin{pmatrix}3\\-2\\4\end{pmatrix}-\begin{pmatrix}\frac{1}{7}\\-\frac{4}{7}\\-\frac{2}{7}\end{pmatrix}=\begin{pmatrix}\frac{20}{7}\\-\frac{10}{7}\\\frac{30}{7}\end{pmatrix}$$

Notemos que $\mathbf{p}\cdot\mathbf{h}=0$.

Concluimos esta sección con un resultado de gran utilidad en estadística y otras áreas de aplicación. Daremos una aplicación de una versión extendida de este teorema en la siguiente sección.

Teorema 8 ***Teorema de aproximación en norma.*** Sea H un subespacio de $\mathbb{R}^n$ y $\mathbf{v}$ un vector de $\mathbb{R}^n$. Entonces en H $\text{proy}_H\mathbf{v}$ es la mejor aproximación a $\mathbf{v}$ en el siguiente sentido: Si $\mathbf{h}$ es cualquier otro vector en H, entonces

$$\|\mathbf{v}-\text{proy}_H\mathbf{v}\| < \|\mathbf{v}-\mathbf{h}\| \qquad (25)$$

Demostración Del Teorema 7, $\mathbf{v}-\text{proy}_H\mathbf{v}\in H^\perp$. Escribimos

$$\mathbf{v}-\mathbf{h} = (\mathbf{v}-\text{proy}_H\mathbf{v}) + (\text{proy}_H\mathbf{v}-\mathbf{h})$$

El primer término en la derecha está en $H^\perp$ y el segundo en H, por lo que

$$(\mathbf{v}-\text{proy}_H\mathbf{v})\cdot(\text{proy}_H\mathbf{v}-\mathbf{h}) = 0 \qquad (26)$$

Ahora,

$$\begin{aligned}\|\mathbf{v}-\mathbf{h}\|^2 &= (\mathbf{v}-\mathbf{h})\cdot(\mathbf{v}-\mathbf{h}) \\ &= [(\mathbf{v}-\text{proy}_H\mathbf{v})+(\text{proy}_H\mathbf{v}-\mathbf{h})]\cdot[(\mathbf{v}-\text{proy}_H\mathbf{v})+(\text{proy}_H\mathbf{v}-\mathbf{h})] \\ &= \|\mathbf{v}-\text{proy}_H\mathbf{v}\|^2 + 2(\mathbf{v}-\text{proy}_H\mathbf{v})\cdot(\text{proy}_H\mathbf{v}-\mathbf{h}) + \|\text{proy}_H\mathbf{v}-\mathbf{h}\|^2 \\ &= \|\mathbf{v}-\text{proy}_H\mathbf{v}\|^2 + \|\text{proy}_H\mathbf{v}-\mathbf{h}\|^2\end{aligned}$$

Pero $\|\text{proj}_H\mathbf{v}-\mathbf{h}\|^2>0$ porque $\mathbf{h}\neq\text{proy}_H\mathbf{v}$. Así que

$$\|\mathbf{v}-\mathbf{h}\|^2 > \|\mathbf{v}-\text{proy}_H\mathbf{v}\|^2$$

o bien

$$\|\mathbf{v}-\mathbf{h}\| > \|\mathbf{v}-\text{proy}_H\mathbf{v}\| \quad ■$$

Problemas 4.10

En los Problemas del 1 al 13 construya una base ortonormal para el espacio vectorial dado o subespacio.

1. En $\mathbb{R}^2$, empezando con los vectores base $\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}-1\\1\end{pmatrix}$
2. $H=\{(x, y)\in\mathbb{R}^2: \; x+y=0\}$
3. $H=\{(x, y)\in\mathbb{R}^2: \; ax+by=0\}$
4. En $\mathbb{R}^2$, empezando con $\begin{pmatrix}a\\b\end{pmatrix}, \begin{pmatrix}c\\d\end{pmatrix}$, en donde $ad-bc\neq 0$.
5. $\pi=\{(x, y, z): \; 2x-y-z=0\}$
6. $\pi=\{(x, y, z): \; 3x-2y+6z=0\}$
7. $L=\{(x, y, z): \; x/2=y/3=z/4\}$
8. $L=\{(x, y, z): \; x=3t, \; y=-2t, \; z=t; \; t \text{ real}\}$
9. $H=\{(x, y, z, w)\in\mathbb{R}^4: \; 2x-y+3z-w=0\}$
10. $\pi=\{(x, y, z): \; ax+by+cz=0\}$, en donde $abc\neq 0$.
11. $L=\{(x, y, z): \; x/a=y/b=z/c\} \; z/c\}$, en donde $abc\neq 0$.
12. $H=\{(x_1, x_2, x_3, x_4, x_5)\in\mathbb{R}^5: \; 2x_1-3x_2+x_3+4x_4-x_5=0\}$
13. H es el espacio solución de

$$\begin{aligned}x-3y+\;z&=0\\-2x+2y-3z&=0\\4x-8y+5z&=0\end{aligned}$$

★ 14. Encuentre una base ortonormal en $\mathbb{R}^4$ que incluya a los vectores

$$\mathbf{u}_1=\begin{pmatrix}1/\sqrt{2}\\0\\1/\sqrt{2}\\0\end{pmatrix} \quad \text{y} \quad \mathbf{u}_2=\begin{pmatrix}-1/2\\1/2\\1/2\\-1/2\end{pmatrix}$$

[*Sugerencia:* Encuentre primero dos vectores $\mathbf{v}_3$ y $\mathbf{v}_4$ para completar la base.]

15. Muestre que $Q=\begin{pmatrix}2/3 & 1/3 & 2/3\\1/3 & 2/3 & -2/3\\-2/3 & 2/3 & 1/3\end{pmatrix}$ es una matriz ortogonal.

16. Muestre que si P y Q son matrices ortogonales de $n \times n$ entonces PQ es ortogonal.
17. Verifique el resultado del Problema 16 con

$$P=\begin{pmatrix}1/\sqrt{2} & -1/\sqrt{2}\\ 1/\sqrt{2} & 1/\sqrt{2}\end{pmatrix} \quad \text{y} \quad Q=\begin{pmatrix}1/3 & -\sqrt{8}/3\\ \sqrt{8}/3 & 1/3\end{pmatrix}$$

18. Muestre que si Q es una matriz ortogonal y simétrica, entonces $Q^2 = I$.
19. Muestre que si Q es ortogonal, entonces $\det Q = \pm 1$.
20. Muestre que para cualquier número real t, la matriz $A = \begin{pmatrix}\operatorname{sen} t & \cos t\\ \cos t & -\operatorname{sen} t\end{pmatrix}$ es ortogonal.
21. Sea $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ un conjunto de vectores en $\mathbb{R}^n$ linealmente independiente. Demuestre que $\mathbf{v}_i \neq \mathbf{0}$ para $i = 1, 2, \ldots, k$. [*Sugerencia:* Si $\mathbf{v}_i = \mathbf{0}$, es fácil encontrar constantes $c_1, c_2, \ldots, c_k$ con $c_i \neq 0$ tales que $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k = \mathbf{0}$.]

En los Problemas del 22 al 28 se dan un subespacio H y un vector $\mathbf{v}$. **(a)** *Encuentre* $\text{proy}_H\mathbf{v}$; **(b)** *encuentre una base ortonormal para $H^\perp$;* **(c)** *escriba* $\mathbf{v}$ *como* $\mathbf{h} + \mathbf{p}$, *en donde* $\mathbf{h} \in H$ *y* $\mathbf{p} \in H^\perp$.

22. $H=\left\{\begin{pmatrix}x\\y\end{pmatrix}\in\mathbb{R}^2:\ x+y=0\right\};\ \mathbf{v}=\begin{pmatrix}-1\\2\end{pmatrix}$
23. $H=\left\{\begin{pmatrix}x\\y\end{pmatrix}\in\mathbb{R}^2:\ ax+by=0\right\};\ \mathbf{v}=\begin{pmatrix}a\\b\end{pmatrix}$
24. $H=\left\{\begin{pmatrix}x\\y\\z\end{pmatrix}\in\mathbb{R}^3:\ ax+by+cz=0\right\};\ \mathbf{v}=\begin{pmatrix}a\\b\\c\end{pmatrix},\ \mathbf{v}\neq\mathbf{0}$
25. $H=\left\{\begin{pmatrix}x\\y\\z\end{pmatrix}\in\mathbb{R}^3:\ 3x-2y+6z=0\right\};\ \mathbf{v}=\begin{pmatrix}-3\\1\\4\end{pmatrix}$
26. $H=\left\{\begin{pmatrix}x\\y\\z\end{pmatrix}\in\mathbb{R}^3:\ x/2=y/3=z/4\right\};\ \mathbf{v}=\begin{pmatrix}1\\1\\1\end{pmatrix}$
27. $H=\left\{\begin{pmatrix}x\\y\\z\\w\end{pmatrix}\in\mathbb{R}^4:\ 2x-y+3z-w=0\right\};\ \mathbf{v}=\begin{pmatrix}1\\-1\\2\\3\end{pmatrix}$
28. $H=\left\{\begin{pmatrix}x\\y\\z\\w\end{pmatrix}\in\mathbb{R}^4:\ x=y \ \text{ y } \ w=3y\right\};\ \mathbf{v}=\begin{pmatrix}-1\\2\\3\\1\end{pmatrix}$
29. Sean $\mathbf{u}_1$ y $\mathbf{u}_2$ dos vectores ortonormales en $\mathbb{R}^n$. Muestre que $|\mathbf{u}_1-\mathbf{u}_2|=\sqrt{2}$.
30. Si $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ son ortonormales, muestre que

$$|\mathbf{u}_1+\mathbf{u}_2+\cdots+\mathbf{u}_n|^2=|\mathbf{u}_1|^2+|\mathbf{u}_2|^2+\cdots+|\mathbf{u}_n|^2=n$$

31. Encuentre una condición para los números a y b tal que $\left\{\begin{pmatrix}a\\b\end{pmatrix}, \begin{pmatrix}b\\-a\end{pmatrix}\right\}$ y $\left\{\begin{pmatrix}a\\b\end{pmatrix}, \begin{pmatrix}-b\\a\end{pmatrix}\right\}$ formen bases ortonormales en $\mathbb{R}^2$.
32. Muestre que *cualquier* base ortonormal en $\mathbb{R}^2$ tiene una de las formas de las bases dadas en el Problema 31.

★ 33. Demuestre la *desigualdad de Cauchy-Schwartz en* $\mathbb{R}^n$: $|\mathbf{u}\cdot\mathbf{v}| \le |\mathbf{u}|\,|\mathbf{v}|$. [*Sugerencia:* Si $\mathbf{u} = \mathbf{0}$ o bien $\mathbf{v} = \mathbf{0}$, el resultado es obvio. Si $\mathbf{u} \neq \mathbf{0}$ y $\mathbf{v} \neq \mathbf{0}$, utilice el hecho de que $|\mathbf{u}/|\mathbf{u}| - \mathbf{v}/|\mathbf{v}|| \ge 0$ y $|\mathbf{u}/|\mathbf{u}| + \mathbf{v}/|\mathbf{v}|| \ge 0$.]
34. Muestre que, en el Problema 33, $|\mathbf{u}\cdot\mathbf{v}| = |\mathbf{u}|\,|\mathbf{v}|$ si y sólo si $\mathbf{u} = \lambda\mathbf{v}$ para algún número real λ.
35. Usando el resultado del Problema 34, demuestre que si $|\mathbf{u} + \mathbf{v}| = |\mathbf{u}| + |\mathbf{v}|$, entonces $\mathbf{u}$ y $\mathbf{v}$ son linealmente dependientes.
36. Usando el resultado del Problema 33, demuestre la *desigualdad del triángulo:* $|\mathbf{u} + \mathbf{v}| \le |\mathbf{u}| + |\mathbf{v}|$. [*Sugerencia:* Desarrolle $|\mathbf{u} + \mathbf{v}|^2$.]

★ 37. Suponga que $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$ son vectores en $\mathbb{R}^n$ (no todos nulos) y que

$$|\mathbf{x}_1+\mathbf{x}_2+\cdots+\mathbf{x}_k|=|\mathbf{x}_1|+|\mathbf{x}_2|+\cdots+|\mathbf{x}_k|$$

Muestre que $\dim \text{gen}\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\} = 1$. [*Sugerencia:* Utilice los resultados de los Problemas 35 y 36.]
38. Sea $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ una base ortonormal en $\mathbb{R}^n$ y sea $\mathbf{v}$ un vector en $\mathbb{R}^n$. Demuestre que $|\mathbf{v}|^2 = |\mathbf{v}\cdot\mathbf{u}_1|^2 + |\mathbf{v}\cdot\mathbf{u}_2|^2 + \cdots + |\mathbf{v}\cdot\mathbf{u}_n|^2$. Esta igualdad se conoce como la *igualdad de Parseval* en $\mathbb{R}^n$.
39. Muestre que para cualquier subespacio H de $\mathbb{R}^n$, $(H^\perp)^\perp = H$.
40. Sean H_1 y H_2 dos subespacios de $\mathbb{R}^n$ y supongamos que $H_1^\perp = H_2^\perp$. Muestre que $H_1 = H_2$.
41. Si H_1 y H_2 son subespacios de $\mathbb{R}^n$, muestre que si $H_1 \subset H_2$, entonces $H_2^\perp \subset H_1^\perp$.
42. Demuestre el teorema de Pitágoras generalizado: Sean $\mathbf{u}$ y $\mathbf{v}$ dos vectores en $\mathbb{R}^n$ con $\mathbf{u} \perp \mathbf{v}$. Entonces

$$\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$$

4.11 Espacios de producto interior y proyecciones

En esta sección se usan propiedades elementales de los números complejos (resumidos en el Apéndice 2); se requiere además, tener cierta familiaridad con el material que se ve en el primer año de Cálculo.

En la Sección 1.5 vimos cómo multiplicar dos vectores en $\mathbb{R}^n$ para obtener un escalar. Tal producto se conoce también como *producto interior*. Antes de presentar una definición general, notemos que en $\mathbb{R}^n$, el producto interior de dos vectores es un número real. En otros aspectos (vea Ejemplo 2 que sigue) el producto interior da un número complejo. Para incluir cualquiera de los casos, en la siguiente definición supondremos que el producto interior de dos vectores es un número complejo; puesto que todo número real es un número complejo, esta definición incluye al caso real también.

Definición 1 **Espacio con producto escalar.** El espacio vectorial complejo V se conoce como un *espacio con producto interior* si para cualquier par de vectores $\mathbf{u}$ y $\mathbf{v}$ en V, existe un único número complejo $(\mathbf{u}, \mathbf{v})$, llamado el *producto interior* de $\mathbf{u}$ y $\mathbf{v}$, tal que si $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ están en V y si $\alpha \in \mathbb{C}$, entonces

i. $(\mathbf{v}, \mathbf{v}) \ge 0$
ii. $(\mathbf{v}, \mathbf{v}) = 0$ si y sólo si $\mathbf{v} = \mathbf{0}$.

iii. $(\mathbf{u}, \mathbf{v}+\mathbf{w}) = (\mathbf{u}, \mathbf{v}) + (\mathbf{u}, \mathbf{w})$
iv. $(\mathbf{u}+\mathbf{v}, \mathbf{w}) = (\mathbf{u}, \mathbf{w}) + (\mathbf{v}, \mathbf{w})$
v. $(\mathbf{u}, \mathbf{v}) = \overline{(\mathbf{v}, \mathbf{u})}$
vi. $(\alpha\mathbf{u}, \mathbf{v}) = \alpha(\mathbf{u}, \mathbf{v})$
vii. $(\mathbf{u}, \alpha\mathbf{v}) = \bar{\alpha}(\mathbf{u}, \mathbf{v})$

La barra en las condiciones (*v*) y (*vii*) denota el conjugado complejo.

Nota. Si $(\mathbf{u}, \mathbf{v})$ es real, entonces $\overline{(\mathbf{u}, \mathbf{v})} = (\mathbf{u}, \mathbf{v})$ y podemos quitar la barra en (*v*).

Ejemplo 1 $\mathbb{R}^n$ es un espacio con producto interior con $(\mathbf{u}, \mathbf{v}) = \mathbf{u} \cdot \mathbf{v}$. Las condiciones (*iii-viii*) están contenidas en el Teorema 1.5.1. Las condiciones (*i*) y (*ii*) se incluyen en el resultado (4.10.9).

Ejemplo 2 En el Ejemplo 4.2.13 definimos el espacio $\mathbb{C}^n$. Sean $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ y $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ dos elementos de $\mathbb{C}^n$. (Recuerde que esto significa que los x_i y los y_i son números complejos.) Ahora definimos

$$(\mathbf{x}, \mathbf{y}) = x_1\bar{y}_1 + x_2\bar{y}_2 + \cdots + x_n\bar{y}_n \tag{1}$$

Para mostrar que la Ecuación (1) define un producto interior necesitamos saber algunas propiedades de los números complejos. Si éstas no son conocidas, consulte el Apéndice 2. Para (*i*):

$$(\mathbf{x}, \mathbf{x}) = x_1\bar{x}_1 + x_2\bar{x}_2 + \cdots + x_n\bar{x}_n = |x_1|^2 + |x_2|^2 + \cdots + |x_n|^2$$

De esta manera se satisfacen (*i*) y (*ii*), puesto que $|x_i|$ es un número real. Las condiciones (*iii*) y (*iv*) se deducen del hecho de que $z_1(z_2 + z_3) = z_1z_2 + z_1z_3$ para los números complejos cualesquiera z_1, z_2 y z_3. La condición (*v*) resulta de que $\overline{z_1z_2} = \bar{z}_1\bar{z}_2$ y $\bar{\bar{z}}_1 = z_1$ de modo que $\overline{x_1\bar{y}_1} = \bar{x}_1y_1$. La condición (*vi*) es obvia. Para (*vii*): $(\mathbf{u}, \alpha\mathbf{v}) = \overline{(\alpha\mathbf{v}, \mathbf{u})} = \overline{(\alpha\mathbf{v}, \bar{\mathbf{u}})} = \bar{\alpha}(\bar{\mathbf{v}}, \bar{\mathbf{u}}) = \bar{\alpha}(\mathbf{u}, \mathbf{v})$. Aquí usamos (*vi*) y (*v*).

Ejemplo 3 En $\mathbb{C}^3$ sea $\mathbf{x} = (1 + i, -3, 4 - 3i)$ y $\mathbf{y} = (2 - i, -i, 2 + i)$. Entonces

$$\begin{aligned}(\mathbf{x}, \mathbf{y}) &= (1+i)\overline{(2-i)} + (-3)\overline{(-i)} + (4-3i)\overline{(2+i)} \\ &= (1+i)(2+i) + (-3)(i) + (4-3i)(2-i) \\ &= (1+3i) - 3i + (5-10i) = 6 - 10i\end{aligned}$$

□ **Ejemplo 4** Supongamos que $a < b$; sea $V = C[a, b]$ y definamos

$$(f, g) = \int_a^b f(t)g(t)\, dt \tag{2}$$

Veamos ahora que esto es un producto interior.

(*i*) $(f, f) = \int_a^b f^2(t)\, dt \geq 0$. Esto se obtiene de un teorema básico de cálculo que establece que si $f \in C[a, b]$, $f \geq 0$ en $[a, b]$, y $\int_a^b f(t)\, dt = 0$, entonces $f = 0$ en $[a, b]$. Esto demuestra (*i*) y (*ii*). Las condiciones (*iii*) a (*vii*) resultan de propiedades elementales de las integrales definidas.

□ **Ejemplo 5** Sea $f(t) = t^2 \in C[0, 1]$ y $g(t) = (4 - t) \in C[0, 1]$. Entonces

$$(f, g) = \int_0^1 t^2(4-t)\, dt = \int_0^1 (4t^2 - t^3)\, dt = \left(\frac{4t^3}{3} - \frac{t^4}{4}\right)\Big|_0^1 = \frac{13}{12}$$

Definición 2 Sea V un espacio con producto interior y supongamos que $\mathbf{u}$ y $\mathbf{v}$ están en V. Entonces:

i. $\mathbf{u}$ y $\mathbf{v}$ son *ortogonales* si $(\mathbf{u}, \mathbf{v}) = 0$.
ii. La *norma* de $\mathbf{u}$, denotada por $|\mathbf{u}|$, está dada por

$$|\mathbf{u}| = \sqrt{(\mathbf{u}, \mathbf{u})} \tag{3}$$

Nota. La Ecuación (3) tiene sentido en virtud de que $(\mathbf{u}, \mathbf{u}) \geq 0$.

Ejemplo 6 En $\mathbb{C}^2$ los vectores $(3, -i)$ y $(2, 6i)$ son ortogonales puesto que

$$((3, -i), (2, 6i)) = 3 \cdot \bar{2} + (-i)\overline{(6i)} = 6 + (-i)(-6i) = 6 - 6 = 0$$

También $|(3, -i)| = \sqrt{3 \cdot 3 + (-i)\overline{(i)}} = \sqrt{10}$.

□ **Ejemplo 7** En $C[0,2\pi]$ las funciones sen t y cos t son ortogonales puesto que

$$(\text{sen}\, t, \cos t) = \int_0^{2\pi} \text{sen}\, t \cos t\, dt = \frac{1}{2}\int_0^{2\pi} \text{sen}\, 2t\, dt = -\frac{\cos 2t}{4}\Big|_0^{2\pi} = 0$$

También

$$\begin{aligned}|\text{sen}\, t| &= (\text{sen}\, t, \text{sen}\, t)^{1/2} \\ &= \left[\int_0^{2\pi} \text{sen}^2\, t\, dt\right]^{1/2} \\ &= \left[\frac{1}{2}\int_0^{2\pi} (1 - \cos 2t)\, dt\right]^{1/2} \\ &= \left[\frac{1}{2}\left(t - \frac{\text{sen}\, 2t}{2}\right)\Big|_0^{2\pi}\right]^{1/2} \\ &= \sqrt{\pi}\end{aligned}$$

Si observamos las demostraciones de los Teoremas 4.10.1 y 4.10.2, veremos que no usamos el hecho de que $V = \mathbb{R}^n$. Estos teoremas son válidos en cualquier espacio con producto interior V, y los enunciaremos después de dar una definición.

Definición 3 **Conjunto ortonormal.** El conjunto de vectores $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ es un *conjunto ortonormal* en V si

$$(\mathbf{v}_i, \mathbf{v}_j) = 0 \quad \text{para} \quad i \neq j \tag{4}$$

y

$$|\mathbf{v}_i| = \sqrt{(\mathbf{v}_i, \mathbf{v}_i)} = 1 \tag{5}$$

Si se cumple (4) solamente, se dice que el conjunto es *ortogonal.*

Teorema 1 Cualquier conjunto ortogonal finito de vectores no nulos en un espacio con un producto interior es linealmente independiente.

Teorema 2 Cualquier conjunto finito y linealmente independiente en un espacio con un producto interior puede transformarse en un conjunto ortonormal por medio del proceso de Gram-Schmidt. En particular, cualquier espacio con producto interior de dimensión finita tiene una base ortonormal.

□ **Ejemplo 8** Construya una base ortonormal para $P_2[0,1]$.

Solución Empezamos con la base estándar $\{1, x, x^2\}$. Dado que $P_2[0, 1]$ es un subespacio de $C[0, 1]$, podemos usar el producto interior del Ejemplo 4. Ya que $\int_0^1 1^2\, dx = 1$, haremos $\mathbf{u}_1 = 1$. Entonces $\mathbf{v}_1' = \mathbf{v}_2 - (\mathbf{v}_2, \mathbf{u}_1)\mathbf{u}_1$. Aquí $(\mathbf{v}_2, \mathbf{u}_1) = \int_0^1 (x \cdot 1)\, dx = \frac{1}{2}$. Por consiguiente, $\mathbf{v}_2' = x - \frac{1}{2} \cdot 1 = x - \frac{1}{2}$. Ahora calculamos

$$|x - \tfrac{1}{2}| = \left[\int_0^1 (x - \tfrac{1}{2})^2\, dx\right]^{1/2} = \left[\int_0^1 (x^2 - x + \tfrac{1}{4})\, dx\right]^{1/2} = \frac{1}{\sqrt{12}} = \frac{1}{2\sqrt{3}}$$

De donde $\mathbf{u}_2 = 2\sqrt{3}(x - \frac{1}{2}) = \sqrt{3}(2x - 1)$. Entonces $\mathbf{v}_3' = \mathbf{v}_3 - (\mathbf{v}_3, \mathbf{u}_1)\mathbf{u}_1 - (\mathbf{v}_3, \mathbf{u}_2)\mathbf{u}_2$. Tenemos $(\mathbf{v}_3, \mathbf{u}_1) = \int_0^1 x^2\, dx = \frac{1}{3}$ y

$$(\mathbf{v}_3, \mathbf{u}_2) = \sqrt{3}\int_0^1 x^2(2x - 1)\, dx = \sqrt{3}\int_0^1 (2x^3 - x^2)\, dx = \frac{\sqrt{3}}{6}$$

Por lo tanto

$$\mathbf{v}_3' = x^2 - \tfrac{1}{3} - \frac{\sqrt{3}}{6}[\sqrt{3}(2x - 1)] = x^2 - x + \tfrac{1}{6}$$

y

$$|\mathbf{v}_3'| = \left[\int_0^1 (x^2 - x + \tfrac{1}{6})^2\, dx\right]^{1/2}$$

$$= \left[\int_0^1 \left(x^4 - 2x^3 + \tfrac{4}{3}x^2 - \frac{x}{3} + \tfrac{1}{36}\right) dx\right]^{1/2}$$

$$= \left[\left(\frac{x^5}{5} - \frac{x^4}{2} + \frac{4x^3}{9} - \frac{x^2}{6} + \frac{x}{36}\right)\Bigg|_0^1\right]^{1/2}$$

$$= \frac{1}{\sqrt{180}} = \frac{1}{6\sqrt{5}}$$

Así, $\mathbf{u}_3 = 6\sqrt{5}(x^2 - x + \frac{1}{6}) = \sqrt{5}(6x^2 - 6x + 1)$. Finalmente, la base ortonormal es $\{1, \sqrt{3}(2x - 1), \sqrt{5}(6x^2 - 6x + 1)\}$.

□ **Ejemplo 9** En $C[0, 2\pi]$, el conjunto infinito

$$S = \left\{\frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}}\operatorname{sen} x, \frac{1}{\sqrt{\pi}}\cos x, \frac{1}{\sqrt{\pi}}\operatorname{sen} 2x, \frac{1}{\sqrt{\pi}}\cos 2x, \ldots, \frac{1}{\sqrt{\pi}}\operatorname{sen} nx, \frac{1}{\sqrt{\pi}}\cos nx, \ldots\right\}$$

es un conjunto ortonormal. Esto resulta de que si $m \neq n$, entonces

$$\int_0^{2\pi} \operatorname{sen} mx \cos nx\, dx = \int_0^{2\pi} \operatorname{sen} mx \operatorname{sen} nx\, dx = \int_0^{2\pi} \cos mx \cos nx\, dx = 0.$$

Para demostrar una de éstas, notamos que

$$\int_0^{2\pi} \operatorname{sen} mx \cos nx\, dx = \frac{1}{2}\int_0^{2\pi} [\operatorname{sen}(m + n)x + \operatorname{sen}(m - n)x]\, dx$$

$$= -\frac{1}{2}\left[\frac{\cos(m + n)x}{m + n} + \frac{\cos(m - n)x}{m - n}\right]\Bigg|_0^{2\pi}$$

$$= 0$$

puesto que $\cos x$ es periódica con periodo 2π. Ya hemos visto que $|\operatorname{sen} x| = \sqrt{\pi}$. Así, $|(1/\sqrt{\pi}) \operatorname{sen} x| = 1$. Los otros hechos se deducen en forma análoga. Este ejemplo nos ilustra una situación en la que tenemos un conjunto ortonormal *infinito*. De hecho, aunque esto está fuera del alcance de este texto, las funciones en S constituyen una base en $C[0, 2\pi]$. Supongamos que $f \in C[0, 2\pi]$. Entonces si escribimos f como una combinación lineal de los vectores en S, obtenemos lo que se conoce como la *representación de f en series de Fourier*.

Definición 4 **Proyección ortogonal.** Sea H un subespacio de un espacio con producto escalar (o interior) V y una base ortonormal $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k\}$. Si $\mathbf{v} \in V$, entonces la *proyección ortogonal* de $\mathbf{v}$ sobre H, se denota por $\text{proy}_H \mathbf{v}$, está dada por

$$\text{proy}_H \mathbf{v} = (\mathbf{v}, \mathbf{u}_1)\mathbf{u}_1 + (\mathbf{v}, \mathbf{u}_2)\mathbf{u}_2 + \cdots + (\mathbf{v}, \mathbf{u}_k)\mathbf{u}_k \tag{6}$$

Las demostraciones de los teoremas siguientes son idénticas a sus contrapartes en $\mathbb{R}^n$ demostradas en la Sección 4.10.

Teorema 3 Sea H un subespacio de un espacio V finito-dimensional con producto interior. Suponga que H tiene dos bases ortonormales $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k\}$ y $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_k\}$. Sea $\mathbf{v} \in V$. Entonces

$$(\mathbf{v}, \mathbf{u}_1)\mathbf{u}_1 + (\mathbf{v}, \mathbf{u}_2)\mathbf{u}_2 + \cdots + (\mathbf{v}, \mathbf{u}_k)\mathbf{u}_k = (\mathbf{v}, \mathbf{w}_1)\mathbf{w}_1 + (\mathbf{v}, \mathbf{w}_2)\mathbf{w}_2 + \cdots + (\mathbf{v}, \mathbf{w}_k)\mathbf{w}_k$$

Definición 5 ***Complemento ortogonal.*** Sea H un subespacio de un espacio V con producto interior. Entonces el complemento ortogonal de H, denotado por $H^\perp$ está dado por

$$H^\perp = \{\mathbf{x} \in V: (\mathbf{x}, \mathbf{h}) = 0 \text{ para toda } \mathbf{h} \in H\} \quad (7)$$

Teorema 4 Si H es un subespacio de un espacio V con producto interior, entonces

i. $H^\perp$ es un subespacio de V.
ii. $H \cap H^\perp = \{\mathbf{0}\}$.
iii. $\dim H^\perp = n - \dim H$ si $\dim V = n < \infty$.

Teorema 5 ***Teorema de la proyección.*** Sea H un subespacio finito-dimensional de un espacio V con producto interior. Entonces, existe un único par de vectores $\mathbf{h}$ y $\mathbf{p}$ tales que $\mathbf{h} \in H$, $\mathbf{p} \in H^\perp$, y

$$\mathbf{v} = \mathbf{h} + \mathbf{p} \quad (8)$$

donde $\mathbf{h} = \text{proy}_H \mathbf{v}$.
Si V es finito-dimensional, $\mathbf{p} = \text{proy}_{H^\perp}\mathbf{v}$.

Observación. Si se examina con cuidado la demostración del Teorema 4.10.7 se advertirá que (8) es válida aun si V es infinito-dimensional. La única diferencia es que, en este caso, $H^\perp$ es de dimensión infinita (pues H es de dimensión finita), y así $\text{proy}_{H^\perp}\mathbf{v}$ no está definida.

Teorema 6 ***Teorema de aproximación en norma.*** Sea H un subespacio finito-dimensional de un espacio V con producto interior, y sea $\mathbf{v}$ un vector en V. Entonces, en H $\text{proy}_H\mathbf{v}$ es la mejor aproximación a $\mathbf{v}$ en el sentido siguiente: Si $\mathbf{h}$ es cualquier otro elemento de H, entonces

$$\|\mathbf{v} - \text{proy}_H\mathbf{v}\| < \|\mathbf{v} - \mathbf{h}\| \quad (9)$$

□ **Ejemplo 10** Como $P_2[0,1]$ es un subespacio de dimensión finita de $C[0,1]$, podemos hablar de $\text{proy}_{P_2[0,1]} f$ si $f \in C[0,1]$. Si $f(x) = e^x$, por ejemplo, calculemos $\text{proy}_{P_2[0,1]} e^x$. Puesto que $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_3\} = \{1, \sqrt{3}(2x-1), \sqrt{5}(6x^2-6x+1)\}$ es una base ortonormal en $P_2[0,1]$ por el Ejemplo 8, tenemos

$$\text{proy}_{P_2[0,1]} e^x = (e^x, 1)1 + (e^x, \sqrt{3}(2x-1))\sqrt{3}(2x-1) + (e^x, \sqrt{5}(6x^2-6x+1))\sqrt{5}(6x^2-6x+1)$$

Teniendo en cuenta que $\int_0^1 e^x dx = e-1$, $\int_0^1 xe^x dx = 1$, y $\int_0^1 x^2e^x dx = e-2$, obtenemos $(e^x, 1) = e-1$, $(e^x, \sqrt{3}(2x-1)) = \sqrt{3}(3-e)$ y $(e^x, \sqrt{5}(6x^2-6x+1)) = \sqrt{5}(7e-19)$. Finalmente,

$$\begin{aligned}\text{proy}_{P_2[0,1]} e^x &= (e-1) + \sqrt{3}(3-e)\sqrt{3}(2x-1) \\ &\quad + \sqrt{5}(7e-19)(\sqrt{5})(6x^2-6x+1) \\ &= (e-1) + (9-3e)(2x-1) \\ &\quad 5(7e-19)(6x^2-6x+1) \\ &\approx 1.01 + 0.85x + 0.84x^2\end{aligned}$$

Concluimos esta sección con una aplicación del teorema de aproximación en norma.

Aproximación cuadrática media de una función continua Sea $f \in C[a,b]$. Se desea aproximar f mediante un polinomio de grado n. ¿Cuál polinomio es el que logra esto con el menor error?

Para poder contestar la pregunta hay que definir lo que entenderemos por *error*. Existen muchos modos distintos de definir error. Damos tres de ellos a continuación:

$$\textbf{error máximo} = \text{máx}\,|f(x) - g(x)| \text{ para } x \in [a,b] \quad (10)$$

$$\textbf{error por área} = \int_a^b |f(x) - g(x)|\,dx \quad (11)$$

$$\textbf{error cuadrático medio} = \int_a^b |f(x) - g(x)|^2\,dx \quad (12)$$

□ **Ejemplo 11** Sea $f(x) = x^2$ y $g(x) = x^3$ en $[0,1]$. Sobre $[0,1]$, $x^2 \geq x^3$ así que $|x^2 - x^3| = x^2 - x^3$. Entonces

i. error máximo $=$ máx $(x^2 - x^3)$. Para calcular este error, se evalúa $d/dx\,(x^2 - x^3) = 2x - 3x^2 = x(2-3x) = 0$ cuando $x = 0$ y $x = 2/3$. El error máximo ocurre cuando $x = 2/3$ y vale $[(2/3)^2 - (2/3)^3] = 4/9 - 8/27 = 4/27 \approx 0.148$.
ii. error por área $= \int_0^1 (x^2 - x^3)dx = (x^3/3 - x^4/4)|_0^1 = 1/3 - 1/4 = 1/12 \approx 0.083$. Esto se representa en la Figura 4.3

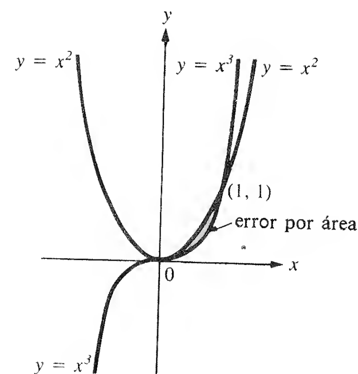


Figura 4.3

iii. error cuadrático medio $= \int_0^1 (x^2 - x^3)^2\,dx = \int_0^1 (x^4 - 2x^5 + x^6)dx = (x^5/5 - x^6/3 + x^7/7)|_0^1 = 1/5 - 1/3 + 1/7 = 1/105 \approx 0.00952$

Cada una de las tres medidas de error es útil. El error cuadrático medio se utiliza en estadística y en otras aplicaciones. Podemos usar el teorema de aproximación en norma para encontrar el polinomio de grado n único que aproxime una función dada con el menor error cuadrático medio.

Del Ejemplo 4, $C[a, b]$ es un espacio con producto interior dado por

$$(f, g) = \int_a^b f(t)g(t)dt \qquad (13)$$

Para todo número entero positivo n, $P_n[a, b]$ —el espacio de polinomios de grado n definidos en $[a, b]$— es un subespacio finito-dimensional de $C[a, b]$. Para $f \in C[a, b]$ y $p_n \in P_n$,

$$\|f - p_n\|^2 = (f - p_n, f - p_n) = \int_a^b [(f(t) - p_n(t))(f(t) - p_n(t)]dt$$

$$= \int_a^b |f(t) - p_n(t)|^2\,dt = \text{error cuadrático medio}$$

Así que por el Teorema 6,

> El polinomio de grado n que aproxima una función continua con el menor error cuadrático medio está dado por $p_n = \text{proy}_{P_n} f$ (14)

Ejemplo 12 Del Ejemplo 10, el polinomio de grado 2 que mejor aproxima la función e^x en [0, 1] en el sentido del error cuadrático medio está dado por

$$p_2(x) \approx 1.01 + 0.85x + 0.84x^2$$

Problemas 4.11

1. Denotemos con D_n al conjunto de matrices diagonales de $n \times n$ con componentes reales, con las operaciones matriciales usuales. Si A y B están en D_n, definamos

$$(A, B) = a_{11}b_{11} + a_{22}b_{22} + \cdots + a_{nn}b_{nn}$$

Demuestre que D_n es un espacio con producto interior.

2. Si $A \in D_n$, demuestre que $|A| = 1$ si y sólo si $a_{11}^2 + a_{22}^2 + \cdots + a_{nn}^2 = 1$.
3. Encuentre una base ortonormal para D_n.
4. Realice la determinación de una base ortonormal para D_2 empezando con $A = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$ y $B = \begin{pmatrix} -3 & 0 \\ 0 & 4 \end{pmatrix}$.
5. En $\mathbb{C}^2$, encuentre una base ortonormal empezando con la base $(1, i)$, $(2 - i, 3 + 2i)$.

□ 6. Encuentre una base ortonormal para $P_3[0, 1]$.

□ 7. Encuentre una base ortonormal para $P_2[-1, 1]$. Los polinomios que obtenga se conocen como *polinomios normalizados de Legendre*.

□★ 8. Encuentre una base ortonormal para $P_2[a, b]$, $a < b$.

9. Si $A = (a_{ij})$ es una matriz $n \times n$, la *traza* de A, denotada con tr A, es la suma de los componentes de la diagonal de A: tr $A = a_{11} + a_{22} + \cdots + a_{nn}$. En M_{nn}, definamos $(A, B) = \text{tr}(AB^t)$. Demuestre que con este producto interior, M_{nn} es un espacio con producto interior.
10. Si $A \in M_{nn}$, muestre que $|A|^2$ es la suma de los cuadrados de los elementos de A. [*Nota:* Aquí $|A| = (A, A)^{1/2}$, no det A.]
11. Encuentre una base ortonormal para M_{22}.
12. Podemos pensar en el plano complejo como un espacio vectorial sobre los reales teniendo por base a los vectores 1, i. Si $z = a + ib$ y $w = c + id$, definamos $(z, w) = ac + bd$. Muestre que esto es un producto interior y que $|z|$ es la magnitud usual de un número complejo.
13. Sean a, b y c tres números reales distintos. Sean p y q en P_2 y definamos $(p, q) = p(a)q(a) + p(b)q(b) + p(c)q(c)$.
    a. Demuestre que (p, q) es un producto interior en P_2.
    b. ¿Es $(p, q) = p(a)q(a) + p(b)q(b)$ un producto interior?
14. En $\mathbb{R}^2$, si $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ y $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$, sea $(\mathbf{x}, \mathbf{y})_* = x_1y_1 + 3x_2y_2$. Muestre que $(x, y)_*$ es un producto interior en $\mathbb{R}^2$.
15. Con el producto interior del Problema 14, calcule $\left|\begin{pmatrix} 2 \\ -3 \end{pmatrix}\right|_*$.
16. En $\mathbb{R}^2$, sea $(\mathbf{x}, \mathbf{y}) = x_1y_1 - x_2y_2$. ¿Es éste un producto interior? ¿Por qué?

★★ 17. Sea V un espacio con producto interior. Demuestre que $|(\mathbf{u}, \mathbf{v})| \le |\mathbf{u}|\,|\mathbf{v}|$. Esta desigualdad se llama la *desigualdad de Cauchy-Schwarz*. [*Sugerencia:* Vea el Problema 4.10.33.]

★ 18. Usando el resultado del Problema 17, demuestre que $|\mathbf{u} + \mathbf{v}| \le |\mathbf{u}| + |\mathbf{v}|$. Esta es la *desigualdad del triángulo*. [*Sugerencia:* Vea el Problema 4.10.36.]

□ 19. En $P_3[0, 1]$ sea H el subespacio generado por $\{1, x^2\}$. Encuentre $H^\perp$.

□★ **20.** En $C[-1, 1]$, sea H el subespacio de funciones pares. Muestre que $H^{\perp}$ está formado por las funciones impares. [*Sugerencia:* f es impar si $f(-x) = -f(x)$ y es par si $f(-x) = f(x)$.]

□★ **21.** $H = P_2[0, 1]$ es un subespacio de $P_3[0, 1]$. Escriba el polinomio $1 + 2x + 3x^2 - x^3$ como $h(x) + p(x)$, donde $h(x) \in H$ y $p(x) \in H^{\perp}$.

★ **22.** Obtenga el polinomio de segundo orden que mejor aproxime a la función sen $\frac{\pi}{2}x$ en el intervalo $[0, 1]$ en el sentido del error cuadrático medio.

★ **23.** Resuelva el Problema 22 para la función $\cos\frac{\pi}{2}x$.

24. Sea A una matriz $m \times n$ con elementos complejos. Entonces la matriz *transpuesta conjugada* de A, denotada por A^*, se define por $(A^*)_{ij} = \overline{a_{ji}}$. Determine A^* si

$$A = \begin{pmatrix} 1 - 2i & 3 + 4i \\ 2i & -6 \end{pmatrix}$$

25. Sea A una matriz $n \times n$ invertible con elementos complejos. A se dice *unitaria* si $A^{-1} = A^*$. Demuestre que la matriz siguiente es unitaria:

$$A = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{2} + \frac{i}{2} \\ \frac{1}{\sqrt{2}} & \frac{1}{2} - \frac{i}{2} \end{pmatrix}$$

★ **26.** Demuestre que una matriz $n \times n$ es unitaria si y sólo si las columnas de A constituyen una base ortonormal de

4.12 Una perspectiva diferente: la existencia de una base

En esta sección demostramos uno de los resultados más importantes del álgebra lineal: *Todo espacio vectorial posee una base*. La demostración es difícil e involucra conceptos que son parte de los fundamentos de la matemática. Será algo laborioso seguir los detalles de la demostración. Sin embargo, una vez que se haya hecho esto, se contará con una apreciación más profunda de una noción matemática fundamental.

Comenzaremos con algunas definiciones.

Definición 1 **Ordenación parcial.** Sea S un conjunto. Una *ordenación parcial* en S, es una relación, denotada por $\le$, la cual está definida para algunos de los pares ordenados de elementos de S y que satisface tres condiciones:

i. $x \le x$ para toda $x \in S$	**ley reflexiva**
ii. Si $x \le y$ y $y \le x$, entonces $x = y$	**ley antisimétrica**
iii. Si $x \le y$ y $y \le z$, entonces $x \le z$	**ley transitiva**

Pudiera darse el caso de que existan elementos x, y en S tales que no se cumpla $x \le y$ ni $y \le x$. Sin embargo, si para todo par $x, y \in S$ se cumple $x \le y$ o bien $y \le x$, entonces la ordenación se llama *ordenación total*. Si $x \le y$ o bien $y \le x$, entonces a los elementos x, y se les llama *comparables*.

Notación. $x < y$ significa $x \le y$ y $x \ne y$.

Ejemplo 1 Los números reales están ordenados parcialmente por $\le$, donde $\le$ significa "menor que o igual a". La ordenación es total.

Ejemplo 2 Sea S un conjunto y $P(S)$, llamado el *conjunto potencia* de S, denota el conjunto de todos los subconjuntos de S.

Se dice de dos subconjuntos A y B de S que $A \le B$ si $A \subseteq B$. La relación de inclusión es una ordenación parcial en $P(S)$. Esto es fácil de demostrar. Tenemos que

i. $A \subseteq A$ para todo subconjunto A.
ii. $A \subseteq B$ y $B \subseteq A$ si y sólo si $A = B$.
iii. Supóngase que $A \subseteq B$ y que $B \subseteq C$. Si $x \in A$, entonces $x \in B$, así que $x \in C$ lo cual significa que $A \subseteq C$.

Como se indica en las circunstancias de la Figura 4.4, la ordenación así definida no es total.

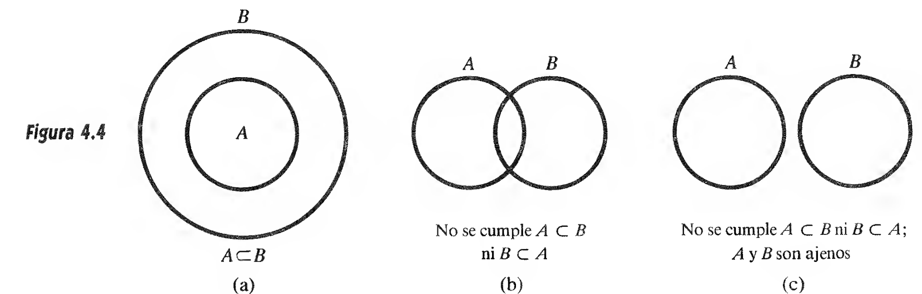


Figura 4.4

Definición 2 ***Cadena, cota superior y elemento maximal.*** Sea S un conjunto parcialmente ordenado por una relación $\le$.

i. Un subconjunto T de S se llama *cadena* si está ordenado totalmente; esto es, si x, y son elementos distintos de T, entonces $x \le y$ o bien $y \le x$.
ii. Sea C un subconjunto de S. Un elemento $u \in S$ se llama *cota superior* de C si $c \le u$ para todo elemento $c \in C$.
iii. El elemento $m \in S$ se denomina *elemento maximal* de S si no existe $s \in S$ siendo $m < s$.

Observación 1. En (*ii*), la cota superior de C debe ser comparable con cada elemento de C pero no necesariamente debe ser elemento de C (aunque sí debe estar en S). Por ejemplo, el número 1 es una cota superior del conjunto (0, 1) pero no pertenece a este intervalo. Todo número mayor que 1 es una cota superior. Sin embargo, no existe número alguno en (0, 1) que sea cota superior de (0, 1).

Observación 2. Si m es un elemento maximal de S, no necesariamente ocurre que $s \le m$ para todo $s \in S$. De hecho, m pudiera ser comparable sólo con algunos

elementos de S. La única condición de maximalidad es que no exista un elemento de S "más grande" que m.

Ejemplo 3 Sea $S = \mathbb{R}^2$. Entonces $P(S)$ consiste en subconjuntos del plano xy. Sea $D_r = \{(x, y): x^2 + y^2 < r^2\}$; esto es, D_r es un disco abierto de radio r —el interior del círculo de radio r centrado en el origen. Sea

$$T = \{D_r: r > 0\}$$

Claramente se ve que T es una cadena, pues si D_{r_1} y D_{r_2} están en T, entonces

$$D_{r_1} \subseteq D_{r_2} \text{ si } r_1 \leq r_2 \quad \text{y} \quad D_{r_2} \subseteq D_{r_1} \text{ si } r_2 \leq r_1$$

Antes de proseguir se necesita una notación adicional. Sea V un espacio vectorial. Hemos visto que una combinación lineal de vectores en V es una suma finita $\sum_{i=1}^{n} \alpha_i \mathbf{v}_i = \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \cdots + \alpha_n \mathbf{v}_n$. Si se han estudiado series de potencias, se habrán visto ya sumas infinitas de la forma $\sum_{n=0}^{\infty} a_n x^n$. Por ejemplo,

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$$

En lo que sigue es necesario un tipo diferente de suma. Sea C un conjunto de vectores en V.* Para cada $\mathbf{v} \in C$, sea α un escalar (el conjunto de escalares se da en la definición de V). Entonces puede escribirse

$$\mathbf{x} = \sum_{\mathbf{v} \in C} \alpha_{\mathbf{v}} \mathbf{v} \qquad (1)$$

se entenderá que sólo un número finito de escalares $\alpha_{\mathbf{v}}$ son distintos de cero y que todos los términos con $\alpha_{\mathbf{v}} = 0$ se eliminan de la suma. Podemos entonces describir la suma (1) como sigue:

> Para cada $\mathbf{v} \in C$, se asigna un escalar $\alpha_{\mathbf{v}}$ y se forma el producto $\alpha_{\mathbf{v}} \mathbf{v}$. Entonces $\mathbf{x}$ es la suma del subconjunto finito de vectores $\alpha_{\mathbf{v}} \mathbf{v}$ para los cuales $\alpha_{\mathbf{v}} \neq 0$.

Definición 3 **Combinación lineal, conjunto generador, independencia lineal y base.**

i. Sea C un subconjunto de un espacio vectorial V. Entonces todo vector que pueda escribirse en la forma (1) se llama *combinación lineal* de vectores en C. El conjunto de combinaciones lineales de vectores en C se denota por $L(C)$.

ii. Del conjunto C se dice que *genera* el espacio vectorial V si $V \subseteq L(C)$.

iii. Un subconjunto C de un espacio vectorial V se dice *linealmente independiente*, si
$$\sum_{\mathbf{v} \in C} \alpha_{\mathbf{v}} \mathbf{v} = \mathbf{0}$$
se verifica sólo si $\alpha_{\mathbf{v}} = 0$ para todo $\mathbf{v} \in C$.

* C no es necesariamente un subespacio de V.

iv. El subconjunto B de un espacio vectorial V es una *base* de V si genera a V y si es linealmente independiente.

Observación. Si C contiene sólo un número finito de vectores, entonces estas definiciones son precisamente las mismas que se vieron con anterioridad en este capítulo.

Teorema 1 Sea B un subconjunto linealmente independiente de un espacio vectorial V. Entonces B es una base si y sólo si es maximal; esto es, si $B \subsetneq D$, entonces D es linealmente dependiente.

Demostración Supóngase que B es una base y que $B \subsetneq D$. Elíjase $\mathbf{x}$ tal que $\mathbf{x} \in D$ pero $\mathbf{x} \notin B$. Como B es una base, $\mathbf{x}$ puede ser escrito como combinación lineal de vectores en B:

$$\mathbf{x} = \sum_{\mathbf{v} \in B} \alpha_{\mathbf{v}} \mathbf{v}$$

Si para toda $\mathbf{v}$ $\alpha_{\mathbf{v}} = 0$, entonces $\mathbf{x} = \mathbf{0}$ y D es dependiente. De otro modo, si para alguna $\mathbf{v}$, $\alpha_{\mathbf{v}} \neq 0$, la suma

$$\mathbf{x} - \sum_{\mathbf{v} \in B} \alpha_{\mathbf{v}} \mathbf{v} = \mathbf{0}$$

pone de manifiesto que D es dependiente. Así pues, B es maximal.

Recíprocamente, supóngase que B es maximal. Sea $\mathbf{x}$ un vector de V que no esté en B. Sea $D = B \cup \{\mathbf{x}\}$. Entonces D es dependiente (porque B es maximal) y existe una ecuación

$$\sum_{\mathbf{v} \in B} \alpha_{\mathbf{v}} \mathbf{v} + \beta \mathbf{x} = \mathbf{0}$$

en la que no todo coeficiente es cero. Pero $\beta \neq 0$, ya que de otro modo obtendríamos una contradicción a la independencia lineal de B. Así pues, puede escribirse

$$\mathbf{x} = -\beta^{-1} \sum_{\mathbf{v} \in B} \alpha_{\mathbf{v}} \mathbf{v}^*$$

De modo que B es un conjunto generador, y por lo tanto, una base de V. ■

¿A dónde lleva todo esto? Quizá el lector ya adivinó la dirección general. Hemos definido ordenaciones en conjuntos y elementos maximales. Se ha demostrado que un conjunto linealmente independiente es una base si es maximal. Hace falta ahora un resultado que ayude en la demostración de la existencia de un elemento maximal. El resultado que buscamos es uno de los supuestos básicos de las matemáticas.

Suele estudiarse la geometría euclidiana en la enseñanza secundaria y preparatoria. Ahí quizá se tuvo un primer contacto con demostraciones matemáticas. Para demostrar cosas, Euclides hizo ciertas suposiciones a las que llamó *axiomas*. Por ejemplo, consideraba que la distancia más corta entre dos puntos es la recta. Comenzando con estos axiomas, él, y los estudiantes de geometría, fueron capaces de demostrar un buen número de teoremas.

* Si los escalares son números reales o complejos, entonces $\beta^{-1} = 1/\beta$.

En todas las ramas de las matemáticas es necesario establecer algunos axiomas. Si nada se supone, nada puede demostrarse. Para completar nuestra demostración, necesitaremos del axioma siguiente:

Axioma **Lema de Zorn.*** Si S es un conjunto no vacío, parcialmente ordenado y tal que toda cadena no vacía posee una cota superior, entonces en S hay un elemento maximal.

Observación. El *axioma de elección* dice, aproximadamente, que dado un número (finito o infinito) de conjuntos no vacíos, existe una función que selecciona un elemento de cada uno de ellos. Este axioma es equivalente al lema de Zorn; esto es, si se postula el axioma de elección, es posible demostrar el lema de Zorn, y viceversa. Si el lector desea ver una demostración de esta equivalencia, así como otros resultados interesantes, se le recomienda el excelente libro *Naive Set Theory* de Paul R. Halmos (Nueva York: Van Nostrand, 1960), especialmente la página 63.

Finalmente, podemos enunciar y demostrar nuestro resultado principal.

Teorema 2 Todo espacio vectorial V tiene una base.

Demostración Se probará que V tiene un subconjunto maximal linealmente independiente. Lo haremos por pasos.

i. Sea S la colección de todos los subconjuntos de V linealmente independientes, y considerémoslos parcialmente ordenados por inclusión.

ii. Una cadena en S es un subconjunto T de S tal que si A y B están en T, entonces $A \subseteq B$ o bien $B \subseteq A$.

iii. Sea T una cadena y definamos

$$M(T) = \bigcup_{A \in T} A$$

Claramente $M(T)$ es un subconjunto de V tal que $A \subseteq M(T)$ para toda $A \in T$. Desea demostrarse que $M(T)$ es una cota superior de T. Como $A \subseteq M(T)$ para toda $A \in T$, sólo hay que probar que $M(T)$ está en S. Esto es, que $M(T)$ es linealmente independiente.

iv. Supóngase que $\sum_{\mathbf{v} \in M(T)} \alpha_\mathbf{v} \mathbf{v} = \mathbf{0}$, donde sólo un número finito de las $\alpha_\mathbf{v}$ son distintas de cero. Denotemos estos escalares por $\alpha_1, \alpha_2, \ldots, \alpha_n$ y sus correspondientes vectores por $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$. Para cada i, $i = 1, 2, \ldots, n$ existe un conjunto $A_i \in T$ tal que $\mathbf{v}_i \in A$ (ya que cada $\mathbf{v}_i$ está en $M(T)$ y $M(T)$ es la unión de conjuntos en T). Pero T está totalmente ordenado, así que una de las A_i contiene a todas las demás (vea Problema 3). Sea este elemento A_k. (Podemos concluir lo anterior porque $\{A_1, A_2, \ldots, A_n\}$ es finito.) Así pues, $A_i \subseteq A_k$ para $i = 1, 2, \ldots, n$ y $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \in A_k$. Como A_k es linealmente independiente y $\sum_{i=1}^{n} \alpha_i \mathbf{v}_i = \mathbf{0}$, se sigue que $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$. Así pues, $M(T)$ es linealmente independiente.

* Max A. Zorn (n. 1906) pasó un buen número de años en la Universidad de Indiana, donde es ahora Profesor Emérito. Publicó su resultado famoso en 1935 ("A Remark on Method in Transfinite Algebra", *Bulletin of the American Mathematical Society*, **41** (1935): 667-670).

v. S no es vacío porque $\emptyset \in S$ (el conjunto vacío se denota por $\emptyset$). Hemos pues demostrado que toda cadena T en S tiene una cota superior $M(T)$ en S. Por el lema de Zorn, S tiene un elemento maximal. Pero S consiste en todos los subconjuntos de V linealmente independientes. El elemento maximal $B \in S$ es, por lo tanto, un subconjunto de V maximal y linealmente independiente. Por el Teorema 1, B es una base de V. ■

Problemas 4.12

1. Demuestre que todo conjunto linealmente independiente en un espacio vectorial V, puede ser expandido a una base.
2. Demuestre que todo conjunto generador en un espacio vectorial V tiene un subconjunto que es una base.
3. Sean n conjuntos $A_1, A_2, \ldots, A_n$ en una cadena T. Demuestre que uno de ellos contiene a todos los demás. [*Sugerencia:* Como T es una cadena se tiene que $A_1 \subseteq A_2$ o bien $A_2 \subseteq A_1$. Así que el resultado es cierto si $n = 2$. Complete la demostración por inducción matemática.]

Ejercicios de repaso • Capítulo 4

En los Ejercicios del 1 *al* 10 *determine si el conjunto dado es un espacio vectorial. Si es así, encuentre su dimensión. Si es de dimensión finita, encuentre una base para el espacio.*

1. Los vectores (x,y,z) en $\mathbb{R}^3$ que satisfacen $x+2y-z=0$
2. Los vectores (x,y,z) en $\mathbb{R}^3$ que satisfacen $x+2y-z\le 0$.
3. Los vectores (x,y,z,w) en $\mathbb{R}^4$ que satisfacen $x+y+z+w=0$.
4. Los vectores en $\mathbb{R}^3$ que satisfacen $x-2=y+3=z-4$
5. El conjunto de matrices triangulares superiores de $n \times n$ con las operaciones de suma de matrices y multiplicación escalar.
6. El conjunto de polinomios de grado ≤ 5
7. El conjunto de polinomios de grado 5
8. El conjunto de matrices de 3×2, $A = (a_{ij})$, con $a_{12} = 0$, con las operaciones de suma matricial y multiplicación escalar.
9. El conjunto del Ejercicio 8, excepto que $a_{12} = 1$
10. El conjunto $S = \{f \in C[0, 2]: f(2) = 0\}$

En los Ejercicios del 11 *al* 19 *determine si el conjunto de vectores dado es linealmente dependiente o independiente.*

11. $\begin{pmatrix}2\\3\end{pmatrix}; \begin{pmatrix}4\\-6\end{pmatrix}$
12. $\begin{pmatrix}2\\3\end{pmatrix}; \begin{pmatrix}4\\6\end{pmatrix}$
13. $\begin{pmatrix}1\\-1\\2\end{pmatrix}; \begin{pmatrix}3\\0\\1\end{pmatrix}; \begin{pmatrix}0\\0\\0\end{pmatrix}$
14. $\begin{pmatrix}1\\-4\\2\end{pmatrix}; \begin{pmatrix}0\\2\\-1\end{pmatrix}; \begin{pmatrix}2\\-10\\5\end{pmatrix}$
15. $\begin{pmatrix}1\\0\\0\\0\end{pmatrix}; \begin{pmatrix}0\\1\\0\\0\end{pmatrix}; \begin{pmatrix}0\\0\\1\\0\end{pmatrix}; \begin{pmatrix}0\\0\\0\\1\end{pmatrix}$
16. En P_3: $1, 2-x^2, 3-x, 7x^2-8x$
17. En P_3: $1, 2+x^3, 3-x, 7x^2-8x$

18. En M_{22}: $\begin{pmatrix}1 & -1\\0 & 0\end{pmatrix}, \begin{pmatrix}1 & 1\\0 & 0\end{pmatrix}, \begin{pmatrix}0 & 0\\1 & 1\end{pmatrix}, \begin{pmatrix}0 & 0\\1 & -1\end{pmatrix}$

19. En M_{22}: $\begin{pmatrix}1 & 1\\0 & 0\end{pmatrix}, \begin{pmatrix}1 & -1\\0 & 0\end{pmatrix}, \begin{pmatrix}0 & 0\\1 & 1\end{pmatrix}, \begin{pmatrix}0 & 0\\1 & -1\end{pmatrix}$

20. Utilizando determinantes, determine si cada conjunto de vectores es independiente o dependiente.

a. $\begin{pmatrix}1\\5\\2\end{pmatrix}; \begin{pmatrix}3\\0\\4\end{pmatrix}; \begin{pmatrix}-5\\5\\6\end{pmatrix}$ b. $(2, 1, 4); (3, -2, 6); (-1, -4, -2)$

En los Ejercicios del 21 *al* 26 *encuentre una base para el espacio vectorial dado y determine su dimensión.*

21. Los vectores en $\mathbb{R}^3$ que están sobre el plano $2x + 3y - 4z = 0$.

22. $H = \{(x, y)\colon\ 2x - 3y = 0\}$

23. $\{\mathbf{v} \in \mathbb{R}^4\colon\ 3x - y - z + w = 0\}$

24. $\{p \in P_3\colon\ p(0) = 0\}$

25. El conjunto de matrices diagonales de 4×4.

26. M_{32}

En los Ejercicios del 27 *al* 32 *determine el núcleo o kernel, la imagen, la nulidad y el rango de la matriz dada.*

27. $A = \begin{pmatrix}1 & -2\\-2 & 4\end{pmatrix}$

28. $A = \begin{pmatrix}1 & -1 & 3\\2 & 0 & 4\\0 & -2 & 2\end{pmatrix}$

29. $A = \begin{pmatrix}1 & -1 & 2\\0 & 1 & 4\\1 & -1 & 0\end{pmatrix}$

30. $A = \begin{pmatrix}2 & 4 & -2\\-1 & -2 & 1\end{pmatrix}$

31. $A = \begin{pmatrix}2 & 3\\-1 & 2\\4 & 6\end{pmatrix}$

32. $A = \begin{pmatrix}1 & -1 & 2 & 3\\0 & 1 & -1 & 0\\1 & -2 & 3 & 3\\2 & -3 & 5 & 6\end{pmatrix}$

En los Ejercicios del 33 *al* 36 *escriba el vector dado en términos de los vectores base.*

33. En $\mathbb{R}^2$: $\mathbf{x} = \begin{pmatrix}2\\-1\end{pmatrix}; \begin{pmatrix}1\\2\end{pmatrix}, \begin{pmatrix}-1\\2\end{pmatrix}$

34. En $\mathbb{R}^3$: $\mathbf{x} = \begin{pmatrix}-3\\4\\2\end{pmatrix}; \begin{pmatrix}1\\0\\1\end{pmatrix}, \begin{pmatrix}1\\1\\0\end{pmatrix}, \begin{pmatrix}0\\2\\3\end{pmatrix}$

35. En P_2: $\mathbf{x} = 4 + x^2;\ 1 + x^2,\ 1 + x,\ 1$

36. En M_{22}: $\mathbf{x} = \begin{pmatrix}3 & 1\\0 & 1\end{pmatrix}; \begin{pmatrix}1 & 1\\0 & 0\end{pmatrix}; \begin{pmatrix}1 & -1\\0 & 0\end{pmatrix}; \begin{pmatrix}0 & 0\\1 & 1\end{pmatrix}; \begin{pmatrix}0 & 0\\1 & -1\end{pmatrix}$

En los Ejercicios 37 *y* 38 *determine el rango de la matriz indicada utilizando subdeterminantes.*

37. $\begin{pmatrix}1 & 3 & -1 & 4\\2 & 1 & 6 & 2\\1 & -2 & 7 & -2\end{pmatrix}$

38. $\begin{pmatrix}2 & 3 & -1\\4 & 1 & 6\\0 & -5 & 8\\-2 & -3 & 3\end{pmatrix}$

En los Ejercicios del 39 *al* 42 *encuentre una base ortonormal para el espacio vectorial dado.*

39. $\mathbb{R}^2$ empezando con la base $\begin{pmatrix}2\\3\end{pmatrix}, \begin{pmatrix}-1\\4\end{pmatrix}$

40. $\{(x, y, z) \in \mathbb{R}^3\colon\ x - y - z = 0\}$

41. $\{(x, y, z) \in \mathbb{R}^3\colon\ x = y = z\}$

42. $\{(x, y, z, w) \in \mathbb{R}^4\colon\ x = z \text{ y } y = w\}$

En los Ejercicios del 43 *al* 45*:* **(a)** calcule $\text{proy}_H \mathbf{v}$; **(b)** *halle una base ortonormal para* $H^\perp$; **(c)** *exprese* $\mathbf{v}$ *como* $\mathbf{h} + \mathbf{p}$, *en donde* $\mathbf{h} \in H$ *y* $\mathbf{p} \in H^\perp$.

43. H es el subespacio del Problema 40; $\mathbf{v} = \begin{pmatrix}-1\\2\\4\end{pmatrix}$.

44. H es el subespacio del Problema 41; $\mathbf{v} = \begin{pmatrix}1\\0\\-1\end{pmatrix}$.

45. H es el subespacio del Problema 42; $\mathbf{v} = \begin{pmatrix}1\\0\\0\\1\end{pmatrix}$.

46. Obtenga una base ortonormal para $P_2[0, 2]$.

47. Utilice el resultado del Ejercicio 46 para encontrar el polinomio que dé la mejor aproximación cuadrática media a e^x en el intervalo $[0, 2]$.

CAPÍTULO 5

Transformaciones lineales

5.1 Definición y ejemplos

En este capítulo vamos a discutir una clase especial de funciones, llamadas *transformaciones lineales*, las cuales aparecen con gran frecuencia en el álgebra lineal y otras ramas de las matemáticas. También son importantes en una amplia variedad de aplicaciones. Antes de definir lo que es una transformación lineal, estudiemos dos ejemplos sencillos para ver lo que puede suceder.

Ejemplo 1 En $\mathbb{R}^2$, definamos una función T por la fórmula $T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ -y \end{pmatrix}$. Geométricamente, T toma un vector en $\mathbb{R}^2$ y lo transforma en su reflexión con respecto al eje x. Esto se ilustra en la Figura 5.1. Una vez que hayamos dado la definición básica, veremos que T es una transformación lineal de $\mathbb{R}^2$ en $\mathbb{R}^2$.

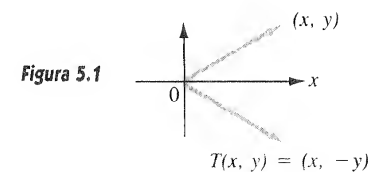


Figura 5.1

Ejemplo 2 Un fabricante elabora cuatro productos diferentes, cada uno de los cuales requiere tres materias primas. Denotamos los cuatro productos por P_1, P_2, P_3 y P_4 y las materias primas por R_1, R_2 y R_3. La tabla anexa da el número de unidades de cada materia prima que se necesitan para la elaboración de una unidad de cada producto.

Necesarias para producir una unidad de

Cantidad de unidades de materia prima		P_1	P_2	P_3	P_4
	R_1	2	1	3	4
	R_2	4	2	2	1
	R_3	3	3	1	2

La pregunta que surge naturalmente es: Si se produce un número determinado de cada uno de los cuatro productos ¿cuántas unidades de cada materia prima se necesitan? Denotemos por p_1, p_2, p_3 y p_4 el número de artículos elaborados de los cuatro productos y por r_1, r_2 y r_3 el número de unidades necesarias de materia prima. Entonces, definimos

$$\mathbf{p}=\begin{pmatrix}p_1\\p_2\\p_3\\p_4\end{pmatrix}\qquad \mathbf{r}=\begin{pmatrix}r_1\\r_2\\r_3\end{pmatrix}\qquad A=\begin{pmatrix}2&1&3&4\\4&2&2&1\\3&3&1&2\end{pmatrix}$$

Por ejemplo, supongamos que $\mathbf{p}=\begin{pmatrix}10\\30\\20\\50\end{pmatrix}$. ¿Cuántas unidades de R_1 se necesitan para producir este número de unidades de los cuatro productos? A partir de la tabla encontramos que

$$\begin{aligned}r_1&=p_1\cdot 2+p_2\cdot 1+p_3\cdot 3+p_4\cdot 4\\&=10\cdot 2+30\cdot 1+20\cdot 3+50\cdot 4=310 \text{ unidades}\end{aligned}$$

Análogamente, $r_2=10\cdot 4+30\cdot 2+20\cdot 2+50\cdot 1=190$ unidades

y $r_3=10\cdot 3+30\cdot 3+20\cdot 1+50\cdot 2=240$ unidades

En general, vemos que

$$\begin{pmatrix}2&1&3&4\\4&2&2&1\\3&3&1&2\end{pmatrix}\begin{pmatrix}p_1\\p_2\\p_3\\p_4\end{pmatrix}=\begin{pmatrix}r_1\\r_2\\r_3\end{pmatrix}$$

o bien

$$\mathbf{r}=A\mathbf{p}$$

Podemos ver esto de otra manera. Si a **p** se le llama *vector de producción*, y a **r**, *vector de materia prima*, definimos la función T por $\mathbf{r}=T\mathbf{p}=A\mathbf{p}$. Esto es, T es la función que "transforma" al vector de producción en el vector de materia prima, que está definido por una multiplicación común de matrices. Como veremos, esta función también es una transformación lineal.

Antes de definir una transformación lineal, conviene expresar algo acerca de las funciones. En la Sección 1.6 escribimos un sistema de ecuaciones como

$$A\mathbf{x}=\mathbf{b}$$

en donde A es una matriz de $m\times n$, $\mathbf{x}\in\mathbb{R}^n$ y $\mathbf{b}\in\mathbb{R}^m$. Se pidió encontrar **x** cuando A y **b** eran conocidos. Sin embargo, es posible considerar a la ecuación de otra manera: Supóngase que A es dada. Entonces la ecuación $A\mathbf{x}=\mathbf{b}$ "dice": Si se da una **x** en $\mathbb{R}^n$, se proporcionará una **b** en $\mathbb{R}^m$. Esto es, A representa una función con dominio en $\mathbb{R}^n$ y un ámbito o contradominio en $\mathbb{R}^m$.

La función definida anteriormente tiene la propiedad de que $A(\alpha\mathbf{x})=\alpha A\mathbf{x}$ si α es un escalar y $A(\mathbf{x}+\mathbf{y})=A\mathbf{x}+A\mathbf{y}$. Esta propiedad caracteriza transformaciones lineales.

Definición 1 ***Transformación lineal.*** Sean V y W espacios vectoriales. Una *transformación lineal* T de V en W es una función que asigna a cada vector $\mathbf{v}\in V$ un único vector $T\mathbf{v}\in W$ y que satisface para cada **u** y **v** en V y cada escalar α,

$$T(\mathbf{u}+\mathbf{v})=T\mathbf{u}+T\mathbf{v} \qquad (1)$$

$$T(\alpha\mathbf{v})=\alpha T\mathbf{v} \qquad (2)$$

Notación Escribimos $T\colon V\to W$ para indicar que T transforma V en W.

Terminología Las transformaciones lineales se llaman, con frecuencia, *operadores lineales.* También, las funciones que satisfacen (1) y (2) se denominan *funciones lineales.*

Ejemplo 3 Sea $T\colon\mathbb{R}^2\to\mathbb{R}^3$ definida por $T\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}x+y\\x-y\\3y\end{pmatrix}$. Por ejemplo, $T\begin{pmatrix}2\\-3\end{pmatrix}=\begin{pmatrix}-1\\5\\-9\end{pmatrix}$. Entonces

$$T\left[\begin{pmatrix}x_1\\y_1\end{pmatrix}+\begin{pmatrix}x_2\\y_2\end{pmatrix}\right]=T\begin{pmatrix}x_1+x_2\\y_1+y_2\end{pmatrix}=\begin{pmatrix}x_1+x_2+y_1+y_2\\x_1+x_2-y_1-y_2\\3y_1+3y_2\end{pmatrix}$$

$$=\begin{pmatrix}x_1+y_1\\x_1-y_1\\3y_1\end{pmatrix}+\begin{pmatrix}x_2+y_2\\x_2-y_2\\3y_2\end{pmatrix}$$

Pero $\begin{pmatrix} x_1+y_1 \\ x_1-y_1 \\ 3y_1 \end{pmatrix} = T\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$ y $\begin{pmatrix} x_2+y_2 \\ x_2-y_2 \\ 3y_2 \end{pmatrix} = T\begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$

Así, $T\left[\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}\right] = T\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + T\begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$

Análogamente,

$$T\left[\alpha\begin{pmatrix} x \\ y \end{pmatrix}\right] = T\begin{pmatrix} \alpha x \\ \alpha y \end{pmatrix} = \begin{pmatrix} \alpha x+\alpha y \\ \alpha x-\alpha y \\ 3\alpha y \end{pmatrix} = \alpha\begin{pmatrix} x+y \\ x-y \\ 3y \end{pmatrix} = \alpha T\begin{pmatrix} x \\ y \end{pmatrix}$$

Por lo tanto T es una transformación lineal.

Ejemplo 4 Sean V y W espacios vectoriales. Defina $T: V \to W$ por $T\mathbf{v} = \mathbf{0}$ para todo $\mathbf{v}$ en V. Entonces $T(\mathbf{v}_1 + \mathbf{v}_2) = \mathbf{0} = \mathbf{0} + \mathbf{0} = T\mathbf{v}_1 + T\mathbf{v}_2$ y $T(\alpha\mathbf{v}) = \mathbf{0} = \alpha\mathbf{0} = \alpha T\mathbf{v}$. Aquí T se llama *transformación cero*.

Ejemplo 5 Sea V un espacio vectorial. Defina $I: V \to V$ por $I\mathbf{v} = \mathbf{v}$ para todo $\mathbf{v}$ en V. Aquí I es obviamente una transformación lineal conocida como *transformación identidad* u *operador identidad*.

Ejemplo 6 Sea $T: \mathbb{R}^2 \to \mathbb{R}^2$ definida por $T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -x \\ y \end{pmatrix}$. Es fácil verificar que T es lineal. Geométricamente T toma un vector en $\mathbb{R}^2$ y lo refleja con respecto al eje y (Figura 5.2).

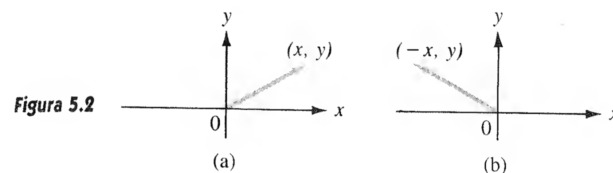


Figura 5.2

Ejemplo 7 Sea A una matriz de $m \times n$. Defina $T: \mathbb{R}^n \to \mathbb{R}^m$ por $T\mathbf{x} = A\mathbf{x}$. Puesto que $A(\mathbf{x}+\mathbf{y}) = A\mathbf{x}+A\mathbf{y}$ y $A(\alpha\mathbf{x}) = \alpha A\mathbf{x}$ si $\mathbf{x}$ y $\mathbf{y}$ están en $\mathbb{R}^n$, vemos que T es una transformación lineal. En consecuencia: *Cada matriz A de $m \times n$ da lugar a una transformación lineal de $\mathbb{R}^n$ en $\mathbb{R}^m$*. En la Sección 5.3 veremos que la recíproca es válida: *Toda transformación lineal entre dos espacios vectoriales de dimensión finita se puede representar por una matriz.*

Ejemplo 8 Suponga que el vector $\mathbf{v} = \begin{pmatrix} x \\ y \end{pmatrix}$ en el plano xy gira un ángulo θ (medido en grados o radianes) en la dirección contraria a las manecillas del reloj.

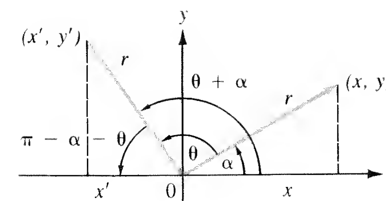


Figura 5.3

Sea $\mathbf{v}' = \begin{pmatrix} x' \\ y' \end{pmatrix}$ el vector girado. Entonces, como en la Figura 5.3, si r denota la longitud de $\mathbf{v}$ (la cual no cambia con la rotación),

$$x = r\cos\alpha \qquad y = r\,\mathrm{sen}\,\alpha$$
$$x' = r\cos(\theta+\alpha) \qquad y' = r\,\mathrm{sen}(\theta+\alpha)^*$$

Pero $r\cos(\theta+\alpha) = r\cos\theta\cos\alpha - r\,\mathrm{sen}\,\theta\,\mathrm{sen}\,\alpha$, por lo que

$$x' = x\cos\theta - y\,\mathrm{sen}\,\theta \qquad (3)$$

Análogamente, $r\,\mathrm{sen}(\theta+\alpha) = r\,\mathrm{sen}\,\theta\cos\alpha + r\cos\theta\,\mathrm{sen}\,\alpha$ o

$$y' = x\,\mathrm{sen}\,\theta + y\cos\theta \qquad (4)$$

Sea

$$A_\theta = \begin{pmatrix} \cos\theta & -\mathrm{sen}\,\theta \\ \mathrm{sen}\,\theta & \cos\theta \end{pmatrix} \qquad (5)$$

Entonces, de (3) y (4), vemos que $A_\theta\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x' \\ y' \end{pmatrix}$. La transformación lineal $T: \mathbb{R}^2 \to \mathbb{R}^2$ definida por $T\mathbf{v} = A_\theta\mathbf{v}$, donde A_θ está dada por (5), se llama *transformación de rotación*.

Ejemplo 9 Sea H un subespacio de $\mathbb{R}^n$. Definimos la *transformación proyección ortogonal* $P: V \to H$ por

$$P\mathbf{v} = \mathrm{proy}_H\,\mathbf{v} \qquad (6)$$

Sea $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k\}$ una base ortonormal de H. Entonces de la Definición 4.10.4 tenemos

$$P\mathbf{v} = (\mathbf{v}\cdot\mathbf{u}_1)\mathbf{u}_1 + (\mathbf{v}\cdot\mathbf{u}_2)\mathbf{u}_2 + \cdots + (\mathbf{v}\cdot\mathbf{u}_k)\mathbf{u}_k \qquad (7)$$

* Esto sale de la definición de $\cos\theta$ y $\mathrm{sen}\,\theta$ como las coordenadas x y y de un punto en el círculo unitario. Si (x, y) es un punto del círculo centrado en el origen con radio r, entonces $x = r\cos\phi$ y $y = r\,\mathrm{sen}\,\phi$, donde ϕ es el ángulo que el vector (x, y) forma con el eje positivo x.

Puesto que $(\mathbf{v}_1+\mathbf{v}_2)\cdot\mathbf{u}=\mathbf{v}_1\cdot\mathbf{u}+\mathbf{v}_2\cdot\mathbf{u}$ y $(\alpha\mathbf{v})\cdot\mathbf{u}=\alpha(\mathbf{v}\cdot\mathbf{u})$, vemos que P es una transformación lineal.

Ejemplo 10 Sea T: $\mathbb{R}^3\to\mathbb{R}^3$ definida por $T\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}x\\y\\0\end{pmatrix}$. Entonces T es el operador proyección que toma un vector en el espacio y lo proyecta en el plano xy. Análogamente $T\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}x\\0\\z\end{pmatrix}$ proyecta un vector del espacio en el plano xz.

Ejemplo 11 Sea T: $M_{mn}\to M_{nm}$ definida por $T(A)=A^t$. Puesto que $(A+B)^t=A^t+B^t$ y $(\alpha A)^t=\alpha A^t$, vemos que T, llamado el *operador transpuesto*, es una transformación lineal.

□ **Ejemplo 12** Sea J: $C[0, 1]\to\mathbb{R}$ y siendo definida por $Jf=\int_0^1 f(x)\,dx$. Puesto que $\int_0^1[f(x)+g(x)]\,dx=\int_0^1 f(x)\,dx+\int_0^1 g(x)\,dx$ y $\int_0^1 \alpha f(x)\,dx=\alpha\int_0^1 f(x)\,dx$ si f y g son continuas, vemos que J es lineal. Por ejemplo, $T(x^3)=\frac{1}{4}$. J se llama *operador integral*.

□ **Ejemplo 13** Sea D: $C'[0, 1]\to C[0, 1]$ definida por $Df=f'$. Puesto que $(f+g)'=f'+g'$ y $(\alpha f)'=\alpha f'$ si f y g son diferenciables, vemos que D es lineal. D se llama *operador diferencial*.

Advertencia. No todas las funciones que tienen apariencia de ser lineales lo son en la realidad. Por ejemplo sea T: $\mathbb{R}\to\mathbb{R}$ definida por $Tx=2x+3$. Entonces $\{(x, Tx): x\in\mathbb{R}\}$ es una línea recta en el plano xy. Pero T no es lineal puesto que $T(x+y)=2(x+y)+3=2x+2y+3$ y $Tx+Ty=(2x+3)+(2y+3)=2x+2y+6$. Las únicas funciones lineales de $\mathbb{R}$ en $\mathbb{R}$ son funciones de la forma $f(x)=mx$ para algún número real m. Por lo tanto, de todas las funciones que son líneas rectas, las únicas que son lineales son las que pasan por el origen.

Ejemplo 14 Sea T: $C[0, 1]\to\mathbb{R}$ definida por $Tf=f(0)+1$. Entonces T no es lineal. Para ver esto, calculemos.

$$T[f+g]=(f+g)(0)+1=f(0)+g(0)+1$$

$$Tf+Tg=[f(0)+1]+[g(0)+1]=f(0)+g(0)+2$$

Este es otro ejemplo de una transformación que a simple vista parece ser lineal pero que no lo es.

Problemas 5.1

En los Problemas del 1 al 29 determine si la transformación dada de V en W es lineal.

1. T: $\mathbb{R}^2\to\mathbb{R}^2$; $T\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}x\\0\end{pmatrix}$
2. T: $\mathbb{R}^2\to\mathbb{R}^2$; $T\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}1\\y\end{pmatrix}$
3. T: $\mathbb{R}^3\to\mathbb{R}^2$; $T\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}x\\y\end{pmatrix}$
4. T: $\mathbb{R}^3\to\mathbb{R}^2$; $T\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}0\\y\end{pmatrix}$
5. T: $\mathbb{R}^3\to\mathbb{R}^2$; $T\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}1\\z\end{pmatrix}$
6. T: $\mathbb{R}^2\to\mathbb{R}^2$; $T\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}x^2\\y^2\end{pmatrix}$
7. T: $\mathbb{R}^2\to\mathbb{R}^2$; $T\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}y\\x\end{pmatrix}$
8. T: $\mathbb{R}^2\to\mathbb{R}^2$; $T\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}x+y\\x-y\end{pmatrix}$
9. T: $\mathbb{R}^2\to\mathbb{R}$; $T\begin{pmatrix}x\\y\end{pmatrix}=xy$
10. T: $\mathbb{R}^n\to\mathbb{R}$; $T\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}=x_1+x_2+\cdots+x_n$
11. T: $\mathbb{R}\to\mathbb{R}^n$; $T(x)=\begin{pmatrix}x\\x\\\vdots\\x\end{pmatrix}$
12. T: $\mathbb{R}^4\to\mathbb{R}^2$; $T\begin{pmatrix}x\\y\\z\\w\end{pmatrix}=\begin{pmatrix}x+z\\y+w\end{pmatrix}$
13. T: $\mathbb{R}^4\to\mathbb{R}^2$; $T\begin{pmatrix}x\\y\\z\\w\end{pmatrix}=\begin{pmatrix}xz\\yw\end{pmatrix}$
14. T: $M_{nn}\to M_{nn}$; $T(A)=AB$, donde B es una matriz de $n\times n$ dada
15. T: $M_{nn}\to M_{nn}$; $T(A)=A^tA$
16. T: $M_{mn}\to M_{mp}$; $T(A)=AB$, donde B es una matriz de $n\times p$ dada
17. T: $D_n\to D_n$; $T(D)=D^2$ (D_n es el conjunto de matrices diagonales de $n\times n$)
18. T: $D_n\to D_n$; $T(D)=I+D$
19. T: $P_2\to P_1$; $T(a_0+a_1x+a_2x^2)=a_0+a_1x$
20. T: $P_2\to P_1$; $T(a_0+a_1x+a_2x^2)=a_1+a_2x$
21. T: $\mathbb{R}\to P_n$; $T(a)=a+ax+ax^2+\cdots+ax^n$
22. T: $P_2\to P_4$; $T(p(x))=[p(x)]^2$
23. T: $C[0, 1]\to C[0, 1]$; $Tf(x)=f^2(x)$
24. T: $C[0, 1]\to C[0, 1]$; $Tf(x)=f(x)+1$
25. □ T: $C[0, 1]\to\mathbb{R}$; $Tf=\int_0^1 f(x)g(x)\,dx$, donde g es una función dada en $C[0,1]$
26. □ T: $C'[0, 1]\to C[0, 1]$; $Tf=(fg)'$, donde g es una función dada en $C'[0, 1]$
27. T: $C[0, 1]\to C[1, 2]$; $Tf(x)=f(x-1)$
28. T: $C[0, 1]\to\mathbb{R}$; $Tf=f(\frac{1}{2})$
29. T: $M_{nn}\to\mathbb{R}$; $T(A)=\det A$
30. Sea T: $\mathbb{R}^2\to\mathbb{R}^2$ dada por $T(x, y)=(-x, -y)$. Describa T geométricamente.
31. Sea T una transformación lineal de $\mathbb{R}^2\to\mathbb{R}^3$ tal que $T\begin{pmatrix}1\\0\end{pmatrix}=\begin{pmatrix}1\\2\\3\end{pmatrix}$ y $T\begin{pmatrix}0\\1\end{pmatrix}=\begin{pmatrix}-4\\0\\5\end{pmatrix}$.

    Encuentre (a) $T\begin{pmatrix}2\\4\end{pmatrix}$ y (b) $T\begin{pmatrix}-3\\7\end{pmatrix}$.

32. En el Ejemplo 8: (a) Encuentre la matriz de rotación A_θ para $\theta = \pi/6$. (b) ¿Qué le sucede al vector $\begin{pmatrix}-3\\4\end{pmatrix}$ si gira un ángulo de $\pi/6$ en el sentido contrario a las manecillas del reloj?

33. Sea $A_\theta = \begin{pmatrix}\cos\theta & -\operatorname{sen}\theta & 0\\ \operatorname{sen}\theta & \cos\theta & 0\\ 0 & 0 & 1\end{pmatrix}$. Describa geométricamente la transformación lineal T: $\mathbb{R}^3 \to \mathbb{R}^3$ dada por $T\mathbf{x} = A_\theta \mathbf{x}$.

34. Responda a las preguntas del Problema 33 para $A_\theta = \begin{pmatrix}\cos\theta & 0 & -\operatorname{sen}\theta\\ 0 & 1 & 0\\ \operatorname{sen}\theta & 0 & \cos\theta\end{pmatrix}$.

35. Suponga que en un espacio vectorial real V, T satisface $T(\mathbf{x} + \mathbf{y}) = T\mathbf{x} + T\mathbf{y}$ y $T(\alpha\mathbf{x}) = \alpha T\mathbf{x}$ para $\alpha \geq 0$. Muestre que T es lineal.

36. Encuentre una transformación lineal T: $M_{33} \to M_{22}$.

37. Si T es una transformación lineal de V en W, muestre que $T(\mathbf{x} - \mathbf{y}) = T\mathbf{x} - T\mathbf{y}$.

38. Si T es una transformación lineal de V en W, muestre que $T\mathbf{0} = \mathbf{0}$. ¿Son los dos vectores cero el mismo vector?

39. Sea V un espacio producto interior y sea $\mathbf{u}_0 \in V$, un vector dado. Sea T: $V \to \mathbb{R}$ (o $\mathbb{C}$) definida por $T\mathbf{v} = (\mathbf{v}, \mathbf{u}_0)$. Muestre que T es lineal.

★ 40. Muestre que si V es un espacio producto interior complejo y T: $V \to \mathbb{C}$ está definida por $T\mathbf{v} = (\mathbf{u}_0, \mathbf{v})$ para un vector dado $\mathbf{u}_0 \in V$, entonces T no es lineal.

41. Sea V un espacio producto interior con un subespacio de dimensión finita H. Sea $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k\}$ una base de H. Muestre que T: $V \to H$ definida por $T\mathbf{v} = (\mathbf{v}, \mathbf{u}_1)\mathbf{u}_1 + (\mathbf{v}, \mathbf{u}_2)\mathbf{u}_2 + \cdots + (\mathbf{v}, \mathbf{u}_k)\mathbf{u}_k$ es una transformación lineal.

42. Sean V y W dos espacios vectoriales. Sea $L(V, W)$ el espacio de transformaciones lineales de V en W. Si T_1 y T_2 son elementos de $L(V, W)$, defínanse αT_1 y $T_1 + T_2$ por $(\alpha T_1)\mathbf{v} = \alpha(T_1\mathbf{v})$ y $(T_1 + T_2)\mathbf{v} = T_1\mathbf{v} = T_2\mathbf{v}$. Demuestre que $L(V, W)$ es un espacio vectorial.

5.2 Propiedades de las transformaciones lineales: Imagen y kernel (o núcleo)

En esta sección desarrollaremos algunas de las propiedades básicas de las transformaciones lineales.

Teorema 1 Sea $T: V \to W$ una transformación lineal. Entonces para todos los vectores $\mathbf{u}$, $\mathbf{v}$, $\mathbf{v}_1$, $\mathbf{v}_2, \ldots, \mathbf{v}_n$ en V y todos los escalares $\alpha_1, \alpha_2, \ldots, \alpha_n$:

i. $T(\mathbf{0}) = \mathbf{0}$
ii. $T(\mathbf{u} - \mathbf{v}) = T\mathbf{u} - T\mathbf{v}$
iii. $T(\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_n\mathbf{v}_n) = \alpha_1 T\mathbf{v}_1 + \alpha_2 T\mathbf{v}_2 + \cdots + \alpha_n T\mathbf{v}_n$

Nota. En la parte (i) el $\mathbf{0}$ de la izquierda, es el vector cero en V mientras que el $\mathbf{0}$ del lado derecho, es el vector cero en W.

Demostración

i. $T(\mathbf{0}) = T(\mathbf{0} + \mathbf{0}) = T(\mathbf{0}) + T(\mathbf{0})$. Entonces $\mathbf{0} = T(\mathbf{0}) - T(\mathbf{0}) = T(\mathbf{0}) + T(\mathbf{0}) - T(\mathbf{0}) = T(\mathbf{0})$.
ii. $T(\mathbf{u} - \mathbf{v}) = T[\mathbf{u} + (-1)\mathbf{v}] = T\mathbf{u} + T[(-1)\mathbf{v}] = T\mathbf{u} + (-1)T\mathbf{v} = T\mathbf{u} - T\mathbf{v}$.
iii. Demostraremos esta parte por inducción (Apéndice 1). Para $n = 2$, obtenemos $T(\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2) = T(\alpha_1\mathbf{v}_1) + T(\alpha_2\mathbf{v}_2) = \alpha_1 T\mathbf{v}_1 + \alpha_2 T\mathbf{v}_2$. En consecuencia, tenemos que, la ecuación es válida para $n = 2$. Supongamos que es válida para $n = k$ y demostremos que vale para $n = k + 1$: $T(\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_k\mathbf{v}_k + \alpha_{k+1}\mathbf{v}_{k+1}) = T(\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_k\mathbf{v}_k) + T(\alpha_{k+1}\mathbf{v}_{k+1})$, y usando la ecuación de la parte (iii) para $n = k$, esto es igual a $(\alpha_1 T\mathbf{v}_1 + \alpha_2 T\mathbf{v}_2 + \alpha_k T\mathbf{v}_k) + \alpha_{k+1} T\mathbf{v}_{k+1}$, que es lo que queríamos demostrar. Con esto queda completa la demostración. ■

Observación. Note que la parte (ii) del Teorema 1, es un caso especial de la parte (iii).

Un hecho importante de las transformaciones lineales es que están completamente determinadas por lo que le hacen a los vectores base.

Teorema 2 Sea V un espacio vectorial de dimensión finita con base $B = \{\mathbf{v}_1, \mathbf{v}_2 \ldots, \mathbf{v}_n\}$. Sean $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n$ n vectores en W. Suponga que T_1 y T_2 son dos transformaciones lineales de V en W tales que $T_1\mathbf{v}_i = T_2\mathbf{v}_i = \mathbf{w}_i$ para $i = 1, 2, \ldots, n$. Entonces para cualquier vector $\mathbf{v} \in V$, $T_1\mathbf{v} = T_2\mathbf{v}$. Es decir $T_1 = T_2$.

Demostración Puesto que B es una base de V, existe un conjunto único de escalares α_1, $\alpha_2, \ldots, \alpha_n$ tales que $\mathbf{v} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_n\mathbf{v}_n$. Entonces, de la parte (iii) del Teorema 1,

$$T_1\mathbf{v} = T_1(\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_n\mathbf{v}_n) = \alpha_1 T_1\mathbf{v}_1 + \alpha_2 T_1\mathbf{v}_2 + \cdots + \alpha_n T_1\mathbf{v}_n$$
$$= \alpha_1\mathbf{w}_1 + \alpha_2\mathbf{w}_2 + \cdots + \alpha_n\mathbf{w}_n$$

Análogamente,

$$T_2\mathbf{v} = T_2(\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_n\mathbf{v}_n) = \alpha_1 T_2\mathbf{v}_1 + \alpha_2 T_2\mathbf{v}_2 + \cdots + \alpha_n T_2\mathbf{v}_n$$
$$= \alpha_1\mathbf{w}_1 + \alpha_2\mathbf{w}_2 + \cdots + \alpha_n\mathbf{w}_n$$

De esta manera, $T_1\mathbf{v} = T_2\mathbf{v}$. ■

El Teorema 2 nos dice que si $T: V \to W$ y V es de dimensión finita, entonces necesitamos conocer únicamente lo que T hace al vector base en V. Esto determina T completamente. Para ver esto, sea $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ una base en V y sea $\mathbf{v}$ otro vector en V. Entonces, como en la demostración del Teorema 2,

$$T\mathbf{v} = \alpha_1 T\mathbf{v}_1 + \alpha_2 T\mathbf{v}_2 + \cdots + \alpha_n T\mathbf{v}_n$$

De esta manera podemos calcular $T\mathbf{v}$ para cualquier vector $\mathbf{v} \in V$ si conocemos $T\mathbf{v}_1$, $T\mathbf{v}_2$, $\ldots$, $T\mathbf{v}_n$.

Ejemplo 1 Sea T una transformación lineal de $\mathbb{R}^3$ en $\mathbb{R}^2$ y suponga que

$T\begin{pmatrix}1\\0\\0\end{pmatrix}=\begin{pmatrix}2\\3\end{pmatrix}$, $T\begin{pmatrix}0\\1\\0\end{pmatrix}=\begin{pmatrix}-1\\4\end{pmatrix}$ y $T\begin{pmatrix}0\\0\\1\end{pmatrix}=\begin{pmatrix}5\\-3\end{pmatrix}$. Calcule $T\begin{pmatrix}3\\-4\\5\end{pmatrix}$.

Solución Tenemos $\begin{pmatrix}3\\-4\\5\end{pmatrix}=3\begin{pmatrix}1\\0\\0\end{pmatrix}-4\begin{pmatrix}0\\1\\0\end{pmatrix}+5\begin{pmatrix}0\\0\\1\end{pmatrix}$.

De esta manera $T\begin{pmatrix}3\\-4\\5\end{pmatrix}=3T\begin{pmatrix}1\\0\\0\end{pmatrix}-4T\begin{pmatrix}0\\1\\0\end{pmatrix}+5T\begin{pmatrix}0\\0\\1\end{pmatrix}$

$$=3\begin{pmatrix}2\\3\end{pmatrix}-4\begin{pmatrix}-1\\4\end{pmatrix}+5\begin{pmatrix}5\\-3\end{pmatrix}=\begin{pmatrix}6\\9\end{pmatrix}+\begin{pmatrix}4\\-16\end{pmatrix}+\begin{pmatrix}25\\-15\end{pmatrix}=\begin{pmatrix}35\\-22\end{pmatrix}$$

Surge otra pregunta: Si $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n$ son n vectores en W, ¿existe una transformación lineal T tal que $T\mathbf{v}_i=\mathbf{w}_i$ para $i=1, 2, \ldots, n$? La respuesta es sí, como se muestra en el siguiente teorema.

Teorema 3 Sea V un espacio vectorial de dimensión finita con base $B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$. Sea también W un espacio vectorial que contiene a los n vectores $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n$. Entonces existe una única transformación lineal $T: V \to W$ tal que $T\mathbf{v}_i = \mathbf{w}_i$ para $i = 1, 2, \ldots, n$.

Demostración Defina una función T del siguiente modo:

i. $T\mathbf{v}_i=\mathbf{w}_i$

ii. Si $\mathbf{v}=\alpha_1\mathbf{v}_1+\alpha_2\mathbf{v}_2+\cdots+\alpha_n\mathbf{v}_n$, entonces

$$T\mathbf{v}=\alpha_1\mathbf{w}_1+\alpha_2\mathbf{w}_2+\cdots+\alpha_n\mathbf{w}_n \tag{1}$$

Ya que B es una base de V, T está definida para todo $\mathbf{v}\in V$; y puesto que W es un espacio vectorial, $T\mathbf{v}\in W$. De esta manera sólo resta demostrar que T es lineal. Pero esto se deduce directamente de la Ecuación (1). Pues si $\mathbf{u}=\alpha_1\mathbf{v}_1+\alpha_2\mathbf{v}_2+\cdots+\alpha_n\mathbf{v}_n$ y $\mathbf{v}=\beta_1\mathbf{v}_1+\beta_2\mathbf{v}_2+\cdots+\beta_n\mathbf{v}_n$, entonces

$$\begin{aligned}T(\mathbf{u}+\mathbf{v})&=T[(\alpha_1+\beta_1)\mathbf{v}_1+(\alpha_2+\beta_2)\mathbf{v}_2+\cdots+(\alpha_n+\beta_n)\mathbf{v}_n]\\&=(\alpha_1+\beta_1)\mathbf{w}_1+(\alpha_2+\beta_2)\mathbf{w}_2+\cdots+(\alpha_n+\beta_n)\mathbf{w}_n=(\alpha_1\mathbf{w}_1+\alpha_2\mathbf{w}_2\\&\quad+\cdots+\alpha_n\mathbf{w}_n)+(\beta_1\mathbf{w}_1+\beta_2\mathbf{w}_2+\cdots+\beta_n\mathbf{w}_n)=T\mathbf{u}+T\mathbf{v}\end{aligned}$$

Análogamente $T(\alpha\mathbf{v}) = \alpha T\mathbf{v}$, por lo tanto T es lineal. La unicidad de T sale del Teorema 2 y con ello el teorema queda demostrado. ■

Observación. En los Teoremas 2 y 3 los vectores $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n$ no necesitan ser distintos. Más aún, hacemos énfasis en que los teoremas son ciertos si V es cualquier espacio vectorial de dimensión finita, no necesariamente $\mathbb{R}^n$. Nótese también que W no tiene que ser necesariamente de dimensión finita.

Ejemplo 2 Encontrar una transformación lineal de $\mathbb{R}^2$, en el plano

$$W=\left\{\begin{pmatrix}x\\y\\z\end{pmatrix}:\quad 2x-y+3z=0\right\}$$

Solución Del Ejemplo 4.6.3 sabemos que W es un subespacio de dimensión dos de $\mathbb{R}^3$ con vectores base $\mathbf{w}_1=\begin{pmatrix}1\\2\\0\end{pmatrix}$ y $\mathbf{w}_2=\begin{pmatrix}0\\3\\1\end{pmatrix}$. Usando la base canónica en $\mathbb{R}^2$, $\mathbf{v}_1=\begin{pmatrix}1\\0\end{pmatrix}$ y $\mathbf{v}_2=\begin{pmatrix}0\\1\end{pmatrix}$, definimos la transformación lineal T por $T\begin{pmatrix}1\\0\end{pmatrix}=\begin{pmatrix}1\\2\\0\end{pmatrix}$ y $T\begin{pmatrix}0\\1\end{pmatrix}=\begin{pmatrix}0\\3\\1\end{pmatrix}$. Entonces, como muestra la discusión que sigue al Teorema 2, T está completamente determinada. Por ejemplo,

$$T\begin{pmatrix}5\\-7\end{pmatrix}=T\left[5\begin{pmatrix}1\\0\end{pmatrix}-7\begin{pmatrix}0\\1\end{pmatrix}\right]=5T\begin{pmatrix}1\\0\end{pmatrix}-7T\begin{pmatrix}0\\1\end{pmatrix}=5\begin{pmatrix}1\\2\\0\end{pmatrix}-7\begin{pmatrix}0\\3\\1\end{pmatrix}=\begin{pmatrix}5\\-11\\-7\end{pmatrix}$$

Volvamos ahora a dos definiciones importantes en la teoría de las transformaciones lineales.

Definición 1 ***Kernel e imagen de una transformación lineal.*** Sean V y W espacios vectoriales y sea $T: V\to W$ una transformación lineal. Entonces:

i. El *kernel* (o *núcleo*) de T, denotado como $\ker T$, está dado por

$$\ker T=\{\mathbf{v}\in V:\quad T\mathbf{v}=\mathbf{0}\} \tag{2}$$

ii. La *imagen* de T, denotada como $\text{imag}\, T$, está dada por

$$\text{imag}\, T=\{\mathbf{w}\in W:\quad \mathbf{w}=T\mathbf{v} \text{ para alguna } \mathbf{v}\in V\} \tag{3}$$

Observación 1. Note que ker T es no vacío ya que por el Teorema 1, $T(\mathbf{0}) = \mathbf{0}$ de manera que $\mathbf{0} \in \ker T$ para toda transformación lineal T. Será interesante encontrar otros vectores en V que sean "mapeados al cero". De nuevo, nótese que cuando escribimos $T(\mathbf{0}) = \mathbf{0}$, el $\mathbf{0}$ de la izquierda está en V, y el $\mathbf{0}$ de la derecha está en W.

Observación 2. El concepto imag T es simplemente el conjunto de "imágenes" de vectores en V bajo la transformación T. De hecho, si $\mathbf{w} = T\mathbf{v}$, diremos que $\mathbf{w}$ es también la *imagen* de $\mathbf{v}$ bajo T.

Antes de dar ejemplos de núcleos o kerneles, e imágenes demostraremos un teorema que será de mucha utilidad.*

Teorema 4 Si $T: V \to W$ es una transformación lineal, entonces:

i. ker T es un subespacio de V.
ii. imag T es un subespacio de W.

Demostración

i. Sean $\mathbf{u}$ y $\mathbf{v}$ en ker T; entonces $T(\mathbf{u} + \mathbf{v}) = T\mathbf{u} + T\mathbf{v} = \mathbf{0} + \mathbf{0} = \mathbf{0}$ y $T(\alpha\mathbf{u}) = \alpha T\mathbf{u} = \alpha\mathbf{0} = \mathbf{0}$ de modo que $\mathbf{u} + \mathbf{v}$ y $\alpha\mathbf{u}$ están en ker T.
ii. Sean $\mathbf{w}$ y $\mathbf{x}$ en imag T. Entonces $\mathbf{w} = T\mathbf{u}$ y $\mathbf{x} = T\mathbf{v}$ para dos vectores $\mathbf{u}$ y $\mathbf{v}$ en V. Esto significa que $T(\mathbf{u}+\mathbf{v}) = T\mathbf{u} + T\mathbf{v} = \mathbf{w} + \mathbf{x}$ y $T(\alpha\mathbf{u}) = \alpha T\mathbf{u} = \alpha\mathbf{w}$. De esta manera $\mathbf{w} + \mathbf{x}$ y $\alpha\mathbf{w}$ están en imag T. ■

Ejemplo 3 Sea $T\mathbf{v} = \mathbf{0}$ para todo $\mathbf{v} \in V$. (T es la transformación cero.) Entonces ker $T = V$ e imag $T = \{\mathbf{0}\}$.

Ejemplo 4 Sea $T\mathbf{v} = \mathbf{v}$ para todo $\mathbf{v} \in V$. (T es la transformación identidad.) Entonces ker $T = \{\mathbf{0}\}$ e imag $T = V$.

Las transformaciones cero e identidad son dos casos extremos. En el primero, todo está en el núcleo. En el segundo, únicamente el vector cero está en el núcleo o kernel. Los casos intermedios son más interesantes.

Ejemplo 5 Sea $T: \mathbb{R}^3 \to \mathbb{R}^3$ definida por $T\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ 0 \end{pmatrix}$. Esto es (Ejemplo 5.1.10): T es el operador proyección de $\mathbb{R}^3$ en el plano xy. Si $T\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} = \mathbf{0} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$, entonces $x = y = 0$. De esta manera ker $T = \{(x, y, z): x = y = 0\}$ = el eje z imag $T = \{(x, y, z): z = 0\}$ = el plano xy. Note que dim ker $T = 1$ y dim imag $T = 2$.

* (N. de R.T.) Suele adoptarse el neologismo kernel (del inglés *kernel*) para designar también al núcleo. El concepto de imagen se denomina *range* en inglés, que no debe traducirse como "rango".

Definición 2 ***Nulidad y rango de una transformación lineal.*** Si T es una transformación lineal de V en W, entonces definimos:

$$\textit{Nulidad de } T = \nu(T) = \dim \ker T \tag{4}$$

$$\textit{Rango de } T = \rho(T) = \dim \operatorname{imag} T \tag{5}$$

Observación. En la Sección 4.7 definimos el rango y nulidad de una matriz. De acuerdo con el Ejemplo 5.1.7, cada matriz A de $m \times n$, origina una transformación lineal $T: \mathbb{R}^n \to \mathbb{R}^m$ definida por $T\mathbf{x} = A\mathbf{x}$. Evidentemente, ker $T = N_A$, imagen T = imagen A, $\nu(T) = \nu(A)$ y $\rho(T) = \rho(A)$. Así se observa que las definiciones de kernel, imagen, nulidad y rango de una transformación lineal son generalizaciones del mismo concepto aplicado a las matrices.

Ejemplo 6 Sea H un subespacio de $\mathbb{R}^n$ y defina (como en el Ejemplo 5.1.9) $T\mathbf{v} = \text{proy}_H\mathbf{v}$. Claramente imag $T = H$. Del Teorema 4.10.7, podemos escribir cualquier $\mathbf{v} \in V$ como $\mathbf{v} = \mathbf{h} + \mathbf{p} = \text{proy}_H\mathbf{v} + \text{proy}_{H^\perp}\mathbf{v}$. Si $T\mathbf{v} = \mathbf{0}$, entonces $\mathbf{h} = \mathbf{0}$, lo cual significa que $\mathbf{v} = \mathbf{p} \in H^\perp$. De esta manera ker $T = H^\perp$, $\rho(T) = \dim H$, y $\nu(T) = \dim H^\perp = n - \rho(T)$.

Ejemplo 7 Sea $V = M_{mn}$ y se define $T: M_{mn} \to M_{mn}$ por $T(A) = A^t$ (ver Ejemplo 5.1.6). Si $TA = A^t = 0$, entonces A^t es la matriz nula de $n \times m$ tal que A es la matriz nula o cero de $m \times n$. Así, ker $T = \{\mathbf{0}\}$ y, claramente, imag $T = M_{nm}$. Esto significa que $\nu(T) = 0$ y $\rho(T) = nm$.

Ejemplo 8 Sea $T: P_3 \to P_2$ definida por $T(p) = T(a_0 + a_1x + a_2x^2 + a_3x^3) = a_0 + a_1x + a_2x^2$. Entonces si $T(p) = 0$, $a_0 + a_1x + a_2x^2 = 0$ para toda x, lo cual implica que $a_0 = a_1 = a_2 = 0$. De esta manera ker $T = \{p \in P_3: p(x) = a_3x^3\}$ e imag $T = P_2$, $\nu(T) = 1$ y $\rho(T) = 3$.

□ **Ejemplo 9** Sea $V = C[0, 1]$ y J: $C[0, 1] \to \mathbb{R}$ definida por $Jf = \int_0^1 f(x)\,dx$ (Ejemplo 5.1.12). Entonces $\ker J = \{f \in C[0,1]: \int_0^1 f(x)\,dx = 0\}$. Sea α un número real. Entonces la función constante $f(x) = \alpha$ para $x \in [0, 1]$ está en $C[0, 1]$ y $\int_0^1 \alpha\,dx = \alpha$. Ya que esto es verdad para todo número real α, tenemos que $\operatorname{imag} J = \mathbb{R}$.

En la siguiente sección veremos cómo toda transformación lineal de un espacio vectorial de dimensión finita en otro puede representarse por una matriz. Esto permitirá que determinemos el núcleo o kernel y la imagen de cualquier transformación lineal entre dos espacios vectoriales de dimensión finita por medio del núcleo y la imagen de la matriz correspondientes.

Problemas 5.2

En los Problemas del 1 *al* 10 *obtener el kernel o núcleo, la imagen, el rango y la nulidad de las transformaciones lineales dadas.*

1. T: $\mathbb{R}^2 \to \mathbb{R}^2$; $T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ 0 \end{pmatrix}$
2. T: $\mathbb{R}^3 \to \mathbb{R}^2$; $T\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} z \\ y \end{pmatrix}$
3. T: $\mathbb{R}^2 \to \mathbb{R}$; $T\begin{pmatrix} x \\ y \end{pmatrix} = x + y$
4. T: $\mathbb{R}^4 \to \mathbb{R}^2$; $T\begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} x+z \\ y+w \end{pmatrix}$
5. T: $M_{22} \to M_{22}$: $T(A) = AB$, donde $B = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$
6. T: $\mathbb{R} \to P_3$: $T(a) = a + ax + ax^2 + ax^3$
7. ★ T: $M_{nn} \to M_{nn}$: $T(A) = A^t + A$
8. □ T: $C'[0, 1] \to C[0, 1]$; $Tf = f'$
9. T: $C[0, 1] \to \mathbb{R}$; $Tf = f(\frac{1}{2})$
10. T: $\mathbb{R}^2 \to \mathbb{R}^2$: T es una rotación de un ángulo de $\pi/3$
11. Sea T: $V \to W$ una transformación lineal; sea $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ una base para V y suponga que $T\mathbf{v}_i = \mathbf{0}$ para $i = 1, 2, \ldots, n$. Muestre que T es la transformación cero.
12. En el Problema 11 suponga que $W = V$ y $T\mathbf{v}_i = \mathbf{v}_i$ para $i = 1, 2, \ldots, n$. Muestre que T es el operador identidad.
13. Sea T: $V \to \mathbb{R}^3$. Demuestre que imag T es: **(a)** $\{\mathbf{0}\}$, **(b)** una recta que pasa por el origen, **(c)** un plano que pasa por el origen, o **(d)** $\mathbb{R}^3$.
14. Sea T: $\mathbb{R}^3 \to V$. Muestre que ker T es uno de los cuatro espacios enunciados en el Problema 13.
15. Encuentre todas las transformaciones lineales de $\mathbb{R}^2$ en $\mathbb{R}^2$ tales que la recta $y = 0$ se transforma en la recta $x = 0$.
16. Encuentre todas las transformaciones lineales de $\mathbb{R}^2$ en $\mathbb{R}^2$ que transforma la recta $y = ax$ en la recta $y = bx$.
17. Realice la obtención de una transformación lineal T de $\mathbb{R}^3 \to \mathbb{R}^3$ tal que $\ker T = \{(x, y, z): 2x - y + z = 0\}$.

18. Realice la obtención de una transformación lineal T de $\mathbb{R}^3 \to \mathbb{R}^3$ tal que $\operatorname{imag} T = \{(x, y, z): 2x - y + z = 0\}$.
19. Sea T: $M_{nn} \to M_{nn}$ definida por $TA = A - A^t$. Demuestre que ker T = {matrices simétricas de $n \times n$} e imag T = {matrices antisimétricas de $n \times n$}.
20. □★ Sea T: $C'[0, 1] \to C[0, 1]$ definida por $Tf(x) = xf'(x)$. Encuentre el núcleo (o kernel) y la imagen de T.
21. ★ En el Problema 5.1.42 se pidió demostrar que el conjunto de las transformaciones lineales de un espacio vectorial V en un espacio vectorial W, forman un espacio vectorial denotado por $L(V, W)$. Suponiendo que $\dim V = n < \infty$ y que $\dim W = m < \infty$, obtenga $\dim L(V, W)$.
22. Sea H un subespacio de V, en donde $\dim H = k$ y $\dim V = n$. Sea U el subconjunto de $L(V, V)$ que tiene la propiedad tal de que si T está en $L(V, V)$, entonces $T\mathbf{h} = \mathbf{0}$ para toda $\mathbf{h}$ en H.
    a. Demuestre que U es un subespacio de $L(V, V)$.
    b. Determine dim U.
23. ★ Sean S, T elementos de $L(V, V)$ tales que ST es la transformación cero. Demuestre o refute lo siguiente: TS es la transformación cero.

5.3 La representación matricial de una transformación lineal

Si A es una matriz de $m \times n$ y T: $\mathbb{R}^n \to \mathbb{R}^m$ está definida por $T\mathbf{x} = A\mathbf{x}$ entonces, como vimos en el Ejemplo 5.1.7, T es una transformación lineal. Ahora veremos que para *toda* transformación lineal de $\mathbb{R}^n$ en $\mathbb{R}^m$, existe una matriz A de $m \times n$ tal que $T\mathbf{x} = A\mathbf{x}$ para toda $\mathbf{x} \in \mathbb{R}^n$. Este hecho es extremadamente útil. Como vimos en la Observación de la página 285, si $T\mathbf{x} = A\mathbf{x}$, entonces $\ker T = N_A$ e $\operatorname{imag} T = R_A$. Más aún, $\nu(T) = \dim \ker T = \nu(A)$ y $\rho(T) = \dim \operatorname{imag} T = \rho(A)$. De esta manera podemos determinar el núcleo, la imagen, la nulidad y el rango de una transformación lineal de $\mathbb{R}^n \to \mathbb{R}^m$ determinando el núcleo y el espacio imagen de una matriz correspondiente. Más aún, cuando se sabe que $T\mathbf{x} = A\mathbf{x}$, podemos evaluar $T\mathbf{x}$ para cualquier $\mathbf{x}$ en $\mathbb{R}^n$ con una simple multiplicación matricial.

Pero esto no es todo. Como veremos, cualquier transformación lineal entre dos espacios vectoriales de dimensión finita puede ser representada por una matriz.

Teorema 1 Sea T: $\mathbb{R}^n \to \mathbb{R}^m$ una transformación lineal. Entonces existe una única matriz A_T de $m \times n$ tal que

$$T\mathbf{x} = A_T\mathbf{x} \quad \text{para toda } \mathbf{x} \in \mathbb{R}^n \qquad (1)$$

Demostración Sea $\mathbf{w}_1 = T\mathbf{e}_1, \mathbf{w}_2 = T\mathbf{e}_2, \ldots, \mathbf{w}_n = T\mathbf{e}_n$. Sea A_T la matriz cuyas columnas son $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n$. Si

$$\mathbf{w}_i = \begin{pmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{mi} \end{pmatrix} \quad \text{para } i = 1, 2, \ldots, n$$

entonces

$$A_T\mathbf{e}_i = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1i} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2i} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mi} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{mi} \end{pmatrix} = \mathbf{w}_i.$$

i-ésima posición

De esta manera $A_T\mathbf{e}_i = \mathbf{w}_i$ para $i = 1, 2, \ldots, n$. Si $\mathbf{x} \in \mathbb{R}^n$, existe un conjunto único de números reales $c_1, c_2, \ldots, c_n$ tales que $\mathbf{x} = c_1\mathbf{e}_1 + c_2\mathbf{e}_2 + \cdots + c_n\mathbf{e}_n$. Entonces $T\mathbf{x} = c_1T\mathbf{e}_1 + c_2T\mathbf{e}_2 + \cdots + c_nT\mathbf{e}_n = c_1\mathbf{w}_1 + c_2\mathbf{w}_2 + \cdots + c_n\mathbf{w}_n$. Pero $A_T\mathbf{x} = A_T(c_1\mathbf{e}_1 + c_2\mathbf{e}_2 + \cdots + c_n\mathbf{e}_n) = c_1A_T\mathbf{e}_1 + c_2A_T\mathbf{e}_2 + \cdots + c_nA_T\mathbf{e}_n = c_1\mathbf{w}_1 + c_2\mathbf{w}_2 + \cdots + c_n\mathbf{w}_n$. Así que $T\mathbf{x} = A_T\mathbf{x}$. Podemos mostrar ahora que A_T es única. Suponga que $T\mathbf{x} = A_T\mathbf{x}$ y $T\mathbf{x} = B_T\mathbf{x}$ para toda $\mathbf{x} \in \mathbb{R}^n$. Entonces $A_T\mathbf{x} = B_T\mathbf{x}$ o haciendo $C_T = A_T - B_T$, tenemos $C_T\mathbf{x} = \mathbf{0}$ para toda $\mathbf{x} \in \mathbb{R}^n$. En particular, $C_T\mathbf{e}_i = \mathbf{0}$ para $i = 1, 2, \ldots, n$. Pero, como vimos en la primera parte del teorema, $C_T\mathbf{e}_i$ es la i-ésima columna de C_T. De esta manera cada una de las n columnas de C_T es el vector cero de m componentes y $C_T = 0$, la matriz cero de $m \times n$. Esto muestra que $A_T = B_T$ y el teorema está demostrado. ■

Observación 1. En este teorema se supuso que cada vector en $\mathbb{R}^n$ y $\mathbb{R}^m$ está en términos de los vectores base canónicos en esos espacios. Si escogemos otras bases para $\mathbb{R}^n$ y $\mathbb{R}^m$, por supuesto que tendremos una matriz A_T diferente. Vea el caso del Ejemplo 4.9.1 o el Ejemplo 8 que se verá luego.

Observación 2. La demostración del teorema nos muestra que A_T es fácil de obtener si formamos la matriz cuyas columnas son los vectores $T\mathbf{e}_i$.

Definición 1 ***Matriz de transformación.*** La matriz A_T en el Teorema 1, se llama *matriz de transformación* correspondiente a T.

En la Sección 5.2 definimos la imagen, el rango, el núcleo y la nulidad de una transformación lineal. En la Sección 4.7 se definió la imagen, el rango, el kernel o núcleo y la nulidad de una matriz. La demostración del siguiente teorema se deriva fácilmente del Teorema 1 y se deja como ejercicio (vea Problema 36).

Teorema 2 Sea A_T la matriz de la transformación que corresponde a la transformación lineal T. Entonces

i. $\text{imag}\, T = R_{A_T} = C_{A_T}$
ii. $\rho(T) = \rho(A_T)$
iii. $\ker T = N_{A_T}$
iv. $\nu(T) = \nu(A_T)$

Ejemplo 1 Encuentre la matriz de transformación A_T correspondiente a la proyección de un vector en $\mathbb{R}^3$ sobre el plano xy.

Solución Aquí $T\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ 0 \end{pmatrix}$. En particular, $T\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, T\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ y $T\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$.

De esta manera $A_T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$. Note que $A_T\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ 0 \end{pmatrix}$.

Ejemplo 2 Sea T: $\mathbb{R}^3 \to \mathbb{R}^4$ definida por

$$T\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x - y \\ y + z \\ 2x - y - z \\ -x + y + 2z \end{pmatrix}$$

Encuentre A_T, $\ker T$, $\text{imag}\, T$, $\nu(T)$ y $\rho(T)$.

Solución $T\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 2 \\ -1 \end{pmatrix}$, $T\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ -1 \\ 1 \end{pmatrix}$, y $T\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ -1 \\ 2 \end{pmatrix}$. De esta manera

$$A_T = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \\ 2 & -1 & -1 \\ -1 & 1 & 2 \end{pmatrix}.$$ Observe (como una comprobación) que

$$\begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \\ 2 & -1 & -1 \\ -1 & 1 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x - y \\ x + z \\ 2x - y - z \\ -x + y + 2z \end{pmatrix}.$$ Lo siguiente es calcular el kernel o núcleo y la imagen de A.

Reduciendo por renglón obtenemos

$$\begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \\ 2 & -1 & -1 \\ -1 & 1 & 2 \end{pmatrix} \xrightarrow[A_{1,4}(1)]{A_{1,3}(-2)} \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 2 \end{pmatrix}$$

$$\xrightarrow[A_{2,3}(-1)]{A_{2,1}(1)} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & -2 \\ 0 & 0 & 2 \end{pmatrix} \xrightarrow{A_{3,4}(1)} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & -2 \\ 0 & 0 & 0 \end{pmatrix}$$

De esta manera $\rho(A) \overset{\text{ya que } \rho(A) + \nu(A) = 3}{=} 3$ y $\nu(A) = 3 - 3 = 0$. Esto significa que ker T = $\{\mathbf{0}\}$, imag T = gen $\left\{\begin{pmatrix} 1 \\ 0 \\ 2 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \\ 2 \end{pmatrix}\right\}$, $\nu(T)=0$ y $\rho(T)=3$.

Ejemplo 3 Sea T: $\mathbb{R}^3 \to \mathbb{R}^3$ definida por $T\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2x - y + 3z \\ 4x - 2y + 6z \\ -6x + 3y - 9z \end{pmatrix}$. Encuentre A_T, ker T, imag T; $\nu(T)$ y $\rho(T)$.

Solución Puesto que $T\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \\ -6 \end{pmatrix}$, $T\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ -2 \\ 3 \end{pmatrix}$, y $T\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 6 \\ -9 \end{pmatrix}$, tenemos

$$A_T = \begin{pmatrix} 2 & -1 & 3 \\ 4 & -2 & 6 \\ -6 & 3 & -9 \end{pmatrix}.$$

Del Ejemplo 4.7.4 vemos que $\rho(A) \overset{\text{Teorema 2}(ii)}{=} \rho(T) = 1$ e imag T = gen $\left\{\begin{pmatrix} 2 \\ 4 \\ -6 \end{pmatrix}\right\}$. Entonces $\nu(T) = 2$. Para encontrar $N_A \overset{\text{Teorema 2}(iii)}{=}$ ker T reducimos por renglón para resolver el sistema $A\mathbf{x} = \mathbf{0}$: $\left(\begin{array}{ccc|c} 2 & -1 & 3 & 0 \\ 4 & -2 & 6 & 0 \\ -6 & 3 & -9 & 0 \end{array}\right) \xrightarrow[A_{1,3}(3)]{A_{1,2}(-2)} \left(\begin{array}{ccc|c} 2 & -1 & 3 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$.

Esto significa que $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in N_A$ si $2x - y + 3z = 0$ o $y = 2x + 3z$. Primero haciendo $x = 1$, $z = 0$ y entonces $x = 0$, $z = 1$, obtenemos una base para N_A:

$$\ker T = N_A = \text{gen}\left\{\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 3 \\ 1 \end{pmatrix}\right\}.$$

Ejemplo 4 Es fácil verificar que si T es la transformación cero de $\mathbb{R}^n \to \mathbb{R}^m$, entonces A_T es la matriz cero de $m \times n$. Análogamente, si T es la transformación identidad de $\mathbb{R}^n \to \mathbb{R}^n$, entonces $A_T = I_n$.

Ejemplo 5 Vimos en el Ejemplo 5.1.8 que si T es la fución que hace girar cada vector en $\mathbb{R}^2$ un ángulo θ, entonces $A_T = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}$.

Ahora generalizaremos el concepto de representación matricial para espacios vectoriales arbitrarios de dimensión finita.

Teorema 3 Sean V un espacio vectorial real de dimensión n, W un espacio vectorial real de dimensión m y $T: V \to W$ una transformación lineal. Sean $B_1 = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ una base de V y $B_2 = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m\}$ una base de W. Entonces existe una única matriz A_T de $m \times n$ tal que

$$(T\mathbf{x})_{B_2} = A_T(\mathbf{x})_{B_1}. \tag{2}$$

Observación 1. La notación en (2) es la notación de la Sección 4.9. Si $\mathbf{x} \in V = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n$, entonces $(\mathbf{x})_{B_1} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$. Sea $\mathbf{c} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$. Entonces $A_T\mathbf{c}$ es un m-vector que denotamos con $\mathbf{d} = \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_m \end{pmatrix}$. La Ecuación (2) dice que $(T\mathbf{x})_{B_2} = \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_m \end{pmatrix}$. Esto es

$$T\mathbf{x} = d_1\mathbf{w}_1 + d_2\mathbf{w}_2 + \cdots + d_m\mathbf{w}_m$$

Observación 2. Como en el Teorema 1, la unicidad de A_T es relativa a las bases B_1 y B_2. Si cambiamos las bases cambiamos A_T (véanse Ejemplos 8, 9 y 10 y Teorema 5).

Demostración Sean $T\mathbf{v}_1 = \mathbf{y}_1$, $T\mathbf{v}_2 = \mathbf{y}_2$, $\ldots$, $T\mathbf{v}_n = \mathbf{y}_n$. Como $\mathbf{y}_i \in W$, tenemos, para $i = 1, 2, \ldots, n$,

$$\mathbf{y}_i = a_{1i}\mathbf{w}_1 + a_{2i}\mathbf{w}_2 + \cdots + a_{mi}\mathbf{w}_m$$

para algún conjunto (único) de escalares $a_{1i}, a_{2i}, \ldots, a_{mi}$, y se tiene así

$$(\mathbf{y}_1)_{B_2} = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}, (\mathbf{y}_2)_{B_2} = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix}, \ldots, (\mathbf{y}_n)_{B_2} = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}$$

Esto significa, por ejemplo, que $\mathbf{y}_1 = a_{11}\mathbf{w}_1 + a_{21}\mathbf{w}_2 + \cdots + a_{m1}\mathbf{w}_m$. Ahora se define

$$A_T = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

Puesto que

$$(\mathbf{v}_1)_{B_1} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, (\mathbf{v}_2)_{B_1} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, (\mathbf{v}_n)_{B_1} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

tenemos, como en la demostración del Teorema 1,

$$A_T(\mathbf{v}_i)_{B_1} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{vmatrix} \begin{vmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{vmatrix} = \begin{pmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{mi} \end{pmatrix} = (\mathbf{y}_i)_{B_2}$$

(i-ésima posición)

Si $\mathbf{x}$ está en V, entonces

$$\mathbf{x} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n$$

$$(\mathbf{x})_{B_1} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$$

y

$$(A_T(\mathbf{x})_{B_1})_{B_2} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \cdot \\ \cdot \\ \cdot \\ c_n \end{pmatrix}$$

$$= \begin{pmatrix} a_{11}c_1 + a_{12}c_2 + \cdots + a_{1n}c_n \\ a_{21}c_1 + a_{22}c_2 + \cdots + a_{2n}c_n \\ \cdot \\ \cdot \\ a_{m1}c_1 + a_{m2}c_2 + \cdots + a_{mn}c_n \end{pmatrix}$$

$$= c_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \cdot \\ \cdot \\ \cdot \\ a_{m1} \end{pmatrix} + c_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \cdot \\ \cdot \\ \cdot \\ a_{m2} \end{pmatrix} + \cdots + c_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \cdot \\ \cdot \\ \cdot \\ a_{mn} \end{pmatrix}$$

$$= c_1(\mathbf{y}_1)_{B_2} + c_2(\mathbf{y}_2)_{B_2} + \cdots + c_n(\mathbf{y}_n)_{B_2}$$

Análogamente, $T\mathbf{x} = T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n) = c_1T\mathbf{v}_1 + c_2T\mathbf{v}_2 + \cdots + c_nT\mathbf{v}_n = c_1\mathbf{y}_1 + c_2\mathbf{y}_2 + \cdots + c_n\mathbf{y}_n$. De esta manera $(T\mathbf{x})_{B_2} = A_T(\mathbf{x})_{B_1}$. La demostración de la unicidad es exactamente como la demostración de unicidad del Teorema 1. ■

El siguiente resultado, de gran utilidad, se deduce inmediatamente del Teorema 4.7.5 y generaliza el Teorema 2. La demostración se deja como ejercicio (Problema 37).

Teorema 4 Sean V y W espacios vectoriales de dimensión finita con $\dim V = n$. Sea $T: V \to W$ una transformación lineal y sea A_T una representación matricial de T. Entonces

i. $\rho(T) = \rho(A_T)$
ii. $\nu(T) = \nu(A_T)$
iii. $\nu(T) + \rho(T) = n$

Ejemplo 6 Sea $T: P_2 \to P_3$ definida por $(Tp)(x) = xp(x)$. Encuentre A_T y úsela para determinar el núcleo y la imagen de T.

Solución Usando las bases canónicas $B_1 = \{1, x, x^2\}$ en P_2 y $B_2 = \{1, x, x^2, x^3\}$ en P_3, tenemos $(T(1))_{B_2} = (x)_{B_2} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$, $(T(x))_{B_2} = (x^2)_{B_2} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$, y $(T(x^2))_{B_2} =$

$(x^3)_{B_2} = \begin{pmatrix}0\\0\\0\\1\end{pmatrix}$. De esta manera $A_T = \begin{pmatrix}0&0&0\\1&0&0\\0&1&0\\0&0&1\end{pmatrix}$. Evidentemente $\rho(A) = 3$ y una base de R_A es $\left\{\begin{pmatrix}0\\1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\\1\end{pmatrix}\right\}$. Por lo tanto, imag T = gen$\{x, x^2, x^3\}$. Puesto que $\nu(A) = 3 - \rho(A) = 0$, vemos que ker $T = \{\mathbf{0}\}$.

Ejemplo 7 Sea $T:P_3 \to P_2$ definida por $T(a_0 + a_1x + a_2x^2 + a_3x^3) = a_1 + a_2x^2$. Calcule A_T y úsela para encontrar núcleo e imagen de T.

Solución Usando las bases canónicas $B_1 = \{1, x, x^2, x^3\}$ en P_3 y $B_2 = \{1, x, x^2\}$ en P_2, inmediatamente vemos que $(T(1))_{B_2} = \begin{pmatrix}0\\0\\0\end{pmatrix}$, $(T(x))_{B_2} = \begin{pmatrix}1\\0\\0\end{pmatrix}$, $(T(x^2))_{B_2} = \begin{pmatrix}0\\0\\1\end{pmatrix}$ y $(T(x^3))_{B_2} = \begin{pmatrix}0\\0\\0\end{pmatrix}$. De esta manera $A_T = \begin{pmatrix}0&1&0&0\\0&0&0&0\\0&0&1&0\end{pmatrix}$. Evidentemente $\rho(A) = 2$ y una base de R_A es $\left\{\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix}\right\}$ por lo que imag T = gen$\{1, x^2\}$. Entonces $\nu(A) = 4 - 2 = 2$; y si $A_T\begin{pmatrix}a_0\\a_1\\a_2\\a_3\end{pmatrix} = \begin{pmatrix}0\\0\\0\end{pmatrix}$, entonces $a_1 = 0$ y $a_2 = 0$. De aquí que a_0 y a_3 son arbitrarios y una base para N_A es $\left\{\begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\\1\end{pmatrix}\right\}$, por lo que una base para ker T es $\{1, x^3\}$.

En todos los ejemplos de esta sección hemos obtenido la matriz A_T usando la base canónica de cada espacio vectorial. Sin embargo, el Teorema 3 es válido para cualquier base en V y W. El siguiente ejemplo demuestra este hecho.

Ejemplo 8 Sea T: $\mathbb{R}^2 \to \mathbb{R}^2$ definida por $T\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}x+y\\x-y\end{pmatrix}$. Usando las bases $B_1 = B_2 = \left\{\begin{pmatrix}1\\-1\end{pmatrix}, \begin{pmatrix}-3\\2\end{pmatrix}\right\}$, calculamos A_T.

Solución Tenemos $T\begin{pmatrix}1\\-1\end{pmatrix} = \begin{pmatrix}0\\2\end{pmatrix}$ y $T\begin{pmatrix}-3\\2\end{pmatrix} = \begin{pmatrix}-1\\-5\end{pmatrix}$. Puesto que , $\begin{pmatrix}0\\2\end{pmatrix} = -6\begin{pmatrix}1\\-1\end{pmatrix} - 2\begin{pmatrix}-3\\2\end{pmatrix}$, encontramos que $\begin{pmatrix}0\\2\end{pmatrix} = \begin{pmatrix}-6\\-2\end{pmatrix}_{B_2}$. Análogamente $\begin{pmatrix}-1\\-5\end{pmatrix} = 17\begin{pmatrix}1\\-1\end{pmatrix} + 6\begin{pmatrix}-3\\2\end{pmatrix}$, por lo que $\begin{pmatrix}-1\\-5\end{pmatrix} = \begin{pmatrix}17\\6\end{pmatrix}_{B_2}$. De esta manera $A_T = \begin{pmatrix}-6&17\\-2&6\end{pmatrix}$. Para calcular $T\begin{pmatrix}-4\\7\end{pmatrix}$, por ejemplo, primero escribimos $\begin{pmatrix}-4\\7\end{pmatrix} = -13\begin{pmatrix}1\\-1\end{pmatrix} - 3\begin{pmatrix}-3\\2\end{pmatrix} = \begin{pmatrix}-13\\-3\end{pmatrix}_{B_1}$. Entonces $T\begin{pmatrix}-4\\7\end{pmatrix} = T\begin{pmatrix}-13\\-3\end{pmatrix}_{B_1} = \left[A_T\begin{pmatrix}-13\\-3\end{pmatrix}\right]_{B_2} = \left[\begin{pmatrix}-6&17\\-2&6\end{pmatrix}\begin{pmatrix}-13\\-3\end{pmatrix}\right]_{B_2} = \begin{pmatrix}27\\8\end{pmatrix}_{B_2} = 27\begin{pmatrix}1\\-1\end{pmatrix} + 8\begin{pmatrix}-3\\2\end{pmatrix} = \begin{pmatrix}3\\-11\end{pmatrix}$. Note que $T\begin{pmatrix}-4\\7\end{pmatrix} = \begin{pmatrix}-4+7\\-4-7\end{pmatrix} = \begin{pmatrix}3\\-11\end{pmatrix}$, lo cual verifica nuestros cálculos.

Para evitar confusión, siempre calcularemos la matriz A_T con respecto a la base canónica, a menos que se especifique de otra manera.* Si T: $V \to V$ es una transformación lineal y se usa otra base B, entonces nos referimos a A_T como la *matriz de transformación de T con respecto a la base B*. De esta manera, en el ejemplo anterior, $A_T = \begin{pmatrix}-6&17\\-2&6\end{pmatrix}$ es la matriz transformación de T con respecto a la base $\left\{\begin{pmatrix}1\\-1\end{pmatrix}, \begin{pmatrix}-3\\2\end{pmatrix}\right\}$.

Antes de terminar esta sección deberemos responder una pregunta obvia. ¿Por qué molestarnos en usar otra base diferente de la canónica, si los cálculos son más complicados, como en el Ejemplo 8? La respuesta es que, con frecuencia, es útil encontrar una base B^* en $\mathbb{R}^n$ de modo que la matriz de transformación con respecto a B^* sea una matriz diagonal. Las matrices diagonales son muy fáciles de manejar y además, como veremos en el Capítulo 6, existen muchas ventajas al escribir una matriz en forma diagonal.

Ejemplo 9 Sea T: $\mathbb{R}^2 \to \mathbb{R}^2$ definida por $T\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}12x+10y\\-15x-13y\end{pmatrix}$. Encuentre A_T con respecto a las bases $B_1 = B_2 = \left\{\begin{pmatrix}1\\-1\end{pmatrix}, \begin{pmatrix}2\\-3\end{pmatrix}\right\}$.

Solución $T\begin{pmatrix}1\\-1\end{pmatrix} = \begin{pmatrix}2\\-2\end{pmatrix} = 2\begin{pmatrix}1\\-1\end{pmatrix} + 0\begin{pmatrix}2\\-3\end{pmatrix} = \begin{pmatrix}2\\0\end{pmatrix}_{B_2}$ y $T\begin{pmatrix}2\\-3\end{pmatrix} = \begin{pmatrix}-6\\9\end{pmatrix} = 0\begin{pmatrix}1\\-1\end{pmatrix} - 3\begin{pmatrix}2\\-3\end{pmatrix} = \begin{pmatrix}0\\-3\end{pmatrix}_{B_2}$. De esta manera $A_T = \begin{pmatrix}2&0\\0&-3\end{pmatrix}$.

* Esto es, en cualquier espacio donde tengamos definida una base canónica.

Existe otra manera de resolver este problema. Los vectores $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ y $\begin{pmatrix} 2 \\ -3 \end{pmatrix}$ están escritos en términos de la base canónica $S = \left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right\}$. Es decir $\begin{pmatrix} 1 \\ -1 \end{pmatrix} = 1\begin{pmatrix} 1 \\ 0 \end{pmatrix} + (-1)\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ y $\begin{pmatrix} 2 \\ -3 \end{pmatrix} = 2\begin{pmatrix} 1 \\ 0 \end{pmatrix} + (-3)\begin{pmatrix} 0 \\ 1 \end{pmatrix}$. De esta manera la matriz $A = \begin{pmatrix} 1 & 2 \\ -1 & -3 \end{pmatrix}$ es la matriz cuya primera y segunda columnas representan las expansiones de los vectores en B_1 en términos de las bases canónicas. Entonces del procedimiento descrito en la página 255, la matriz $A^{-1} = \begin{pmatrix} 3 & 2 \\ -1 & -1 \end{pmatrix}$ es la matriz de transición de S a B_1. Análogamente, la matriz A es la matriz de transición de B_1 a S (vea Problema 4.9.38). Ahora supongamos que $\mathbf{x}$ está escrito en términos de B_1. Entonces $A\mathbf{x}$ es el mismo vector pero escrito en términos de S. Sea $C = \begin{pmatrix} 12 & 10 \\ -15 & -13 \end{pmatrix}$. Entonces $CA\mathbf{x} = T(A\mathbf{x})$ es la imagen de $A\mathbf{x}$ expresada en términos de S. Finalmente, puesto que queremos $T(A\mathbf{x})$ en términos de B_1 (que eso era el problema), multiplicamos del lado izquierdo para la matriz de transición A^{-1} para obtener $(T\mathbf{x})_{B_1} = (A^{-1}CA)(\mathbf{x})_{B_1}$. Es decir,

$$A_T = A^{-1}CA = \begin{pmatrix} 3 & 2 \\ -1 & -1 \end{pmatrix}\begin{pmatrix} 12 & 10 \\ -15 & -13 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ -1 & -3 \end{pmatrix} = \begin{pmatrix} 3 & 2 \\ -1 & -1 \end{pmatrix}\begin{pmatrix} 2 & -6 \\ -2 & 9 \end{pmatrix}$$
$$= \begin{pmatrix} 2 & 0 \\ 0 & -3 \end{pmatrix}$$

como antes. Resumimos este resultado a continuación.

Teorema 5 Sea T: $\mathbb{R}^n \to \mathbb{R}^m$ una transformación lineal. Suponga que C es la matriz transformación de T con respecto a las bases canónicas S_n y S_m en $\mathbb{R}^n$ y $\mathbb{R}^m$, respectivamente. Sea A_1 la matriz de transición de S_n a la base B_1 en $\mathbb{R}^n$ y sea A_2 la matriz de transición de S_m a la base B_2 en $\mathbb{R}^m$. Si A_T denota la matriz de transformación de T con respecto a las bases B_1 y B_2, entonces

$$A_T = A_2CA_1^{-1} \tag{3}$$

En el Ejemplo 9 vimos que en la transformación lineal T con respecto a una nueva base, la matriz de transformación A_T resultó ser una matriz diagonal. Volveremos a este procedimiento de "diagonalización" en la Sección 6.3. Esto es, dada una transformación lineal de $\mathbb{R}^n$ en $\mathbb{R}^n$, veremos que con frecuencia es posible encontrar una base B tal que la matriz de transformación de T con respecto a B sea diagonal.

Ejemplo 10 Sea T: $\mathbb{R}^3 \to \mathbb{R}^2$ definida por $T\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x + y - z \\ 2x - y + z \end{pmatrix}$. Encuentre A_T con respecto a las bases $B_1 = \left\{\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}\right\}$ y $B_2 = \left\{\begin{pmatrix} 4 \\ 3 \end{pmatrix}, \begin{pmatrix} -1 \\ 5 \end{pmatrix}\right\}$.

Solución Con respecto a las bases canónicas, la representación matricial de T es $C = \begin{pmatrix} 1 & 1 & -1 \\ 2 & -1 & 1 \end{pmatrix}$. Del procedimiento presentado en la página 236, A_1 es la inversa de $\begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 1 & 0 & 1 \end{pmatrix}$; esto es, $A_1^{-1} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 1 & 0 & 1 \end{pmatrix}$. Análogamente, $A_2 = \begin{pmatrix} 4 & -1 \\ 3 & 5 \end{pmatrix}^{-1} = \frac{1}{23}\begin{pmatrix} 5 & 1 \\ -3 & 4 \end{pmatrix}$. Entonces, de la Ecuación (3),

$$A_T = A_2CA_1^{-1} = \frac{1}{23}\begin{pmatrix} 5 & 1 \\ -3 & 4 \end{pmatrix}\begin{pmatrix} 1 & 1 & -1 \\ 2 & -1 & 1 \end{pmatrix}\begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 1 & 0 & 1 \end{pmatrix}$$
$$= \frac{1}{23}\begin{pmatrix} 5 & 1 \\ -3 & 4 \end{pmatrix}\begin{pmatrix} 0 & 0 & -2 \\ 3 & -3 & 2 \end{pmatrix} = \frac{1}{23}\begin{pmatrix} 3 & -3 & -8 \\ 12 & -12 & 14 \end{pmatrix}$$

Por ejemplo, sea $\mathbf{x} = \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix}$. Entonces $T\mathbf{x} = \begin{pmatrix} -1 \\ 4 \end{pmatrix}$ en términos de las bases canónicas. En términos de B_1 vemos que $(\mathbf{x})_{B_1} = \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix}_{B_1} = \frac{7}{2}\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + \frac{5}{2}\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{7}{2} \\ \frac{5}{2} \\ \frac{1}{2} \end{pmatrix}$. De igual manera, $(T\mathbf{x})_{B_2} = \begin{pmatrix} -1 \\ 4 \end{pmatrix}_{B_2} = -\frac{1}{23}\begin{pmatrix} 4 \\ 3 \end{pmatrix} + \frac{19}{23}\begin{pmatrix} -1 \\ 5 \end{pmatrix} = \begin{pmatrix} -\frac{1}{23} \\ \frac{19}{23} \end{pmatrix}$. Finalmente, en términos de las bases B_1 y B_2 verificamos que

$$(T\mathbf{x})_{B_2} = A_T(\mathbf{x})_{B_1} = \frac{1}{23}\begin{pmatrix} 3 & -3 & -8 \\ 12 & -12 & 14 \end{pmatrix}\begin{pmatrix} \frac{7}{2} \\ \frac{5}{2} \\ \frac{1}{2} \end{pmatrix} = \frac{1}{23}\begin{pmatrix} -1 \\ 19 \end{pmatrix} = \begin{pmatrix} -\frac{1}{23} \\ \frac{19}{23} \end{pmatrix}$$

Problemas 5.3

En los Problemas del 1 *al* 30 *encuentre la representación matricial A_T de la transformación lineal T,* ker T, imag T, *$\nu(T)$ y $\rho(T)$. A menos que se especifique de otro modo, supondremos que B_1 y B_2 son bases canónicas.*

1. T: $\mathbb{R}^2 \to \mathbb{R}^2$; $T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x - 2y \\ -x + y \end{pmatrix}$

2. T: $\mathbb{R}^2 \to \mathbb{R}^3$; $T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + y \\ x - y \\ 2x + 3y \end{pmatrix}$

3. T: $\mathbb{R}^3 \to \mathbb{R}^2$; $T\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x - y + z \\ -2x + 2y - 2z \end{pmatrix}$

4. T: $\mathbb{R}^2 \to \mathbb{R}^2$; $T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix}$

5. T: $\mathbb{R}^3 \to \mathbb{R}^3$; $T\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x - y + 2z \\ 3x + y + 4z \\ 5x - y + 8z \end{pmatrix}$

6. T: $\mathbb{R}^3 \to \mathbb{R}^3$; $T\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -x + 2y + z \\ 2x - 4y - 2z \\ -3x + 6y + 3z \end{pmatrix}$

7. T: $\mathbb{R}^4 \to \mathbb{R}^3$; $T\begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} x - y + 2z + 3w \\ y + 4z + 3w \\ x + 6z + 6w \end{pmatrix}$

8. T: $\mathbb{R}^4 \to \mathbb{R}^4$; $T\begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} x - y + 2z + w \\ -x + z + 2w \\ x - 2y + 5z + 4w \\ 2x - y + z - w \end{pmatrix}$

9. T: $\mathbb{R}^2 \to \mathbb{R}^2$; $T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x - 2y \\ 2x + y \end{pmatrix}$; $B_1 = B_2 = \left\{ \begin{pmatrix} 1 \\ -2 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \end{pmatrix} \right\}$

10. T: $\mathbb{R}^2 \to \mathbb{R}^2$; $T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 4x - y \\ 3x + 2y \end{pmatrix}$; $B_1 = B_2 = \left\{ \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 4 \\ 3 \end{pmatrix} \right\}$

11. T: $\mathbb{R}^3 \to \mathbb{R}^2$; $T\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2x + y + z \\ y - 3z \end{pmatrix}$;

$B_1 = \left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}$; $B_2 = \left\{ \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \end{pmatrix} \right\}$

12. T: $\mathbb{R}^2 \to \mathbb{R}^3$; $T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x - y \\ 2x + y \\ y \end{pmatrix}$; $B_1 = \left\{ \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right\}$; $B_2 = \left\{ \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 5 \end{pmatrix} \right\}$

13. T: $P_2 \to P_3$; $T(a_0 + a_1x + a_2x^2) = a_1 - a_1x + a_0x^3$

14. T: $\mathbb{R} \to P_3$; $T(a) = a + ax + ax^2 + ax^3$

15. T: $P_3 \to \mathbb{R}$; $T(a_0 + a_1x + a_2x^2 + a_3x^3) = a_2$

16. T: $P_3 \to P_1$; $T(a_0 + a_1x + a_2x^2 + a_3x^3) = (a_1 + a_3)x - a_2$

17. T: $P_3 \to P_2$; $T(a_0 + a_1x + a_2x^2 + a_3x^3) = (a_0 - a_1 + 2a_2 + 3a_3) + (a_1 + 4a_2 + 3a_3)x + (a_0 + 6a_2 + 5a_3)x^2$

18. T: $M_{22} \to M_{22}$; $T\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a - b + 2c + d & -a + 2c + 2d \\ a - 2b + 5c + 4d & 2a - b + c - d \end{pmatrix}$

19. T: $M_{22} \to M_{22}$; $T\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a + b + c + d & a + b + c \\ a + b & a \end{pmatrix}$

20. T: $P_2 \to P_3$; $T[p(x)] = xp(x)$; $B_1 = \{1, x, x^2\}$; $B_2 = \{1, (1+x), (1+x)^2, (1+x)^3\}$

□ 21. D: $P_4 \to P_3$; $Dp(x) = p'(x)$

□ 22. T: $P_4 \to P_4$; $Tp(x) = xp'(x) - p(x)$

□★ 23. D: $P_n \to P_{n-1}$; $Dp(x) = p'(x)$

□ 24. D: $P_4 \to P_2$; $Dp(x) = p''(x)$

□★ 25. T: $P_4 \to P_4$; $Tp(x) = p''(x) + xp'(x) + 2p(x)$

□★ 26. D: $P_n \to P_{n-k}$; $Dp(x) = p^{(k)}(x)$

□★ 27. T: $P_n \to P_n$; $Tp(x) = x^n p^{(n)}(x) + x^{n-1}p^{(n-1)}(x) + \cdots + xp'(x) + p(x)$

□ 28. J: $P_n \to \mathbb{R}$; $Jp = \int_0^1 p(x)\,dx$

29. T: $\mathbb{R}^3 \to P_2$; $T\begin{pmatrix} a \\ b \\ c \end{pmatrix} = a + bx + cx^2$

30. T: $P_3 \to \mathbb{R}^3$; $T(a_0 + a_1x + a_2x^2 + a_3x^3) = \begin{pmatrix} a_3 - a_2 \\ a_1 + a_3 \\ a_2 - a_1 \end{pmatrix}$

31. Sea T: $M_{mn} \to M_{nm}$ dada por $TA = A^T$. Encuentre A_T con respecto a las bases canónicas en M_{mn} y M_{nm}.

★ 32. Sea T: $\mathbb{C}^2 \to \mathbb{C}^2$ dada por $T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + iy \\ (1+i)y - x \end{pmatrix}$. Encuentre A_T.

□ 33. Sea $V = \text{gen}\,\{1, \text{sen}\,x, \cos x\}$. Encuentre A_D, donde $D: V \to V$ está definida por $Df(x) = f'(x)$. Encuentre imag D y ker D.

□ 34. Responda a las preguntas del Problema 33 dado $V = \text{gen}\,\{e^x, xe^x, x^2e^x\}$.

35. Sea T: $\mathbb{C}^2 \to \mathbb{C}^2$ dada por $T\mathbf{x} = \text{proy}_H\mathbf{x}$, donde $H = \text{gen}\,\{(1/\sqrt{2})(1, i)\}$. Encuentre A_T.

36. Demuestre el Teorema 2.

37. Demuestre el Teorema 4.

5.4 Isomorfismos

En esta sección introducimos una terminología importante y demostramos un teorema que nos dice que todos los vectores espaciales de dimensión n son "esencialmente" los mismos.

Definición 1 **Transformación biunívoca.** Sea $T: V \to W$ una transformación lineal. Entonces T es una *transformación biunívoca* o *uno a uno*, y se simboliza por 1-1, si

$$T\mathbf{v}_1 = T\mathbf{v}_2 \quad \text{implica que} \quad \mathbf{v}_1 = \mathbf{v}_2 \tag{1}$$

Esto es, T es 1-1 si todo vector $\mathbf{w}$ en la imagen de T es la imagen de a lo sumo un vector en V.

Teorema 1 Sea $T: V \to W$ una transformación lineal. Entonces T es 1-1 si y sólo si ker $T = \{\mathbf{0}\}$.

Demostración Suponga que ker $T = \{\mathbf{0}\}$ y $T\mathbf{v}_1 = T\mathbf{v}_2$. Entonces $T\mathbf{v}_1 - T\mathbf{v}_2 = T(\mathbf{v}_1 - \mathbf{v}_2) = \mathbf{0}$, lo cual significa que $(\mathbf{v}_1 - \mathbf{v}_2) \in \ker T = \{\mathbf{0}\}$. De esta manera $\mathbf{v}_1 - \mathbf{v}_2 = \mathbf{0}$, por lo tanto $\mathbf{v}_1 = \mathbf{v}_2$, lo cual demuestra que T es 1 -1. Ahora suponga que T es 1-1 y $\mathbf{v} \in \ker T$. Entonces $T\mathbf{v} = \mathbf{0}$. Pero además $T\mathbf{0} = \mathbf{0}$. De esta manera, puesto que T es 1-1, $\mathbf{v} = \mathbf{0}$. Esto completa la demostración. ■

Ejemplo 1 Sea T: $\mathbb{R}^2 \to \mathbb{R}^2$ definida por $T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x - y \\ 2x + y \end{pmatrix}$. Fácilmente encontramos que $A_T = \begin{pmatrix} 1 & -1 \\ 2 & 1 \end{pmatrix}$ y $\rho(A_T) = 2$; de aquí $\nu(A_T) = 0$ y $N_{A_T} = \ker T = \{\mathbf{0}\}$. De esta manera T es 1-1.

Ejemplo 2 Sea $T: \mathbb{R}^2 \to \mathbb{R}^2$ definida por $T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x - y \\ 2x - 2y \end{pmatrix}$. Entonces $A_T = \begin{pmatrix} 1 & -1 \\ 2 & -2 \end{pmatrix}$, $\rho(A_T) = 1$ y $\nu(A_T) = 1$; de aquí $\nu(T) = 1$ y T no es 1-1. Note, por ejemplo, que $T\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \mathbf{0} = T\begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

Definición 2 ***Transformación sobre.*** Sea $T: V \to W$ una transformación lineal. Entonces T se denomina *transformación sobre* W o, simplemente, *sobre*, si para toda $\mathbf{w} \in W$ existe al menos una $\mathbf{v} \in V$ tal que $T\mathbf{v} = \mathbf{w}$. Es decir: *T es sobre W si y sólo si imag $T = W$.*

Ejemplo 3 En el Ejemplo 1, $\rho(A_T) = 2$; de aquí imag $T = \mathbb{R}^2$ y T es sobre. En el Ejemplo 2, $\rho(A_T) = 1$ e imag $T = \text{gen}\left\{\begin{pmatrix} 1 \\ 2 \end{pmatrix}\right\} \neq \mathbb{R}^2$; de aquí que T no es sobre.

Teorema 2 Sea $T: V \to W$ una transformación lineal y suponga que $\dim V = \dim W = n$.

i. Si T es 1-1, entonces T es sobre.
ii. Si T es sobre, entonces T es 1-1.

Demostración Sea A_T la representación matricial de T. Entonces si T es 1-1, $\ker T = \{\mathbf{0}\}$ y $\nu(A_T) = 0$, lo cual significa que $\rho(T) = \rho(A_T) = n - 0 = n$ por lo que imag $T = W$. Si T es sobre, entonces $\rho(A_T) = n$ por lo que $\nu(T) = \nu(A_T) = 0$ y T es 1-1. ■

Teorema 3 Sea $T: V \to W$ una transformación lineal. Suponga que $\dim V = n$ y que $\dim W = m$. Entonces:

i. Si $n > m$, T no es 1-1.
ii. Si $m > n$, T no es sobre.

Demostración i. Sea $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ una base para V. Sea $\mathbf{w}_i = T\mathbf{v}_i$ para $i = 1, 2, \ldots, n$ y observe el conjunto $S = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n\}$. Puesto que $m = \dim W < n$, el conjunto S es linealmente dependiente. De esta manera, existen escalares no todos ceros tales que $c_1\mathbf{w}_1 + c_2\mathbf{w}_2 + \cdots + c_n\mathbf{w}_n = \mathbf{0}$. Sea $\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n$. Puesto que las $\mathbf{v}_i$ son linealmente independientes, y puesto que no todas las c_i son cero, vemos que $\mathbf{v} \neq \mathbf{0}$. Pero $T\mathbf{v} = T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n) = c_1T\mathbf{v}_1 + c_2T\mathbf{v}_2 + \cdots + c_nT\mathbf{v}_n = c_1\mathbf{w}_1 + c_2\mathbf{w}_2 + \cdots + c_n\mathbf{w}_n = \mathbf{0}$. De esta manera $\mathbf{v} \in \ker T$ y $\ker T \neq \{\mathbf{0}\}$.

ii. Si $\mathbf{v} \in V$, entonces $\mathbf{v} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_n\mathbf{v}_n$ para algunos escalares $a_1, a_2, \ldots, a_n$ y $T\mathbf{v} = a_1T\mathbf{v}_1 + a_2T\mathbf{v}_2 + \cdots + a_nT\mathbf{v}_n = a_1\mathbf{w}_1 + a_2\mathbf{w}_2 + \cdots + a_n\mathbf{w}_n$. De esta manera $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n\} = \{T\mathbf{v}_1, T\mathbf{v}_2, \ldots, T\mathbf{v}_n\}$ genera imag T. Entonces, del Problema 4.6.29, $\rho(T) = \dim \text{imag}\, T \leq n$. Puesto que $m > n$, esto muestra que imag $T \neq W$. De esta manera T no es sobre. ■

Definición 3 ***Isomorfismo.*** Sea $T: V \to W$ una transformación lineal. Entonces T es un *isomorfismo* si T es 1-1 y sobre.

Definición 4 ***Espacios vectoriales isomórficos.*** Se dice que los espacios vectoriales V y W son *isomorfos* si existe un isomorfismo T de V sobre W. En este caso escribimos $V \cong W$.

Observación. La palabra "isomorfismo" proviene del griego *isomorphos* que significa "de igual forma" (*iso* = igual; *morphos* = forma). Después de unos cuantos ejemplos veremos cuán estrechamente están relacionadas las "formas" de espacios vectoriales isomorfos.

Sea $T: \mathbb{R}^n \to \mathbb{R}^n$ y sea A_T la representación matricial de T. Sabemos que T es 1-1 si y sólo si $\ker T = \{\mathbf{0}\}$, lo cual es verdad si y sólo si $\nu(A_T) = 0$ si y sólo si $\det A_T \neq 0$. De esta manera podemos ampliar el Teorema Resumen (visto por última vez en la página 225) en otra dirección.

Teorema 4 ***Teorema resumen, Versión 7.*** Sea A una matriz de $n \times n$. Entonces los diez siguientes enunciados son equivalentes. Es decir, si uno es válido, todos son válidos.

i. A es invertible.
ii. La única solución para el sistema homogéneo $A\mathbf{x} = \mathbf{0}$ es la solución trivial ($\mathbf{x} = \mathbf{0}$).
iii. El sistema $A\mathbf{x} = \mathbf{b}$ tiene una única solución para todo n-vector $\mathbf{b}$.
iv. A es equivalente por renglón a la matriz identidad I_n de $n \times n$.
v. A puede ser expresada como el producto de matrices elementales.
vi. Los renglones (y las columnas) de A son linealmente independientes.
vii. $\det A \neq 0$.
viii. $\nu(A) = 0$.
ix. $\rho(A) = n$.
x. La transformación lineal T de $\mathbb{R}^n$ a $\mathbb{R}^n$ definida por $T\mathbf{x} = A\mathbf{x}$ es un isomorfismo.

Veamos ahora algunos ejemplos de isomorfismos entre otros pares de espacios vectoriales.

Ejemplo 4 Sea $T: \mathbb{R}^3 \to P_2$ definida por $T\begin{pmatrix} a \\ b \\ c \end{pmatrix} = a + bx + cx^2$. Es fácil verificar que T es

lineal. Suponga que $T\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \mathbf{0} = 0 + 0x + 0x^2$. Entonces $a = b = c = 0$. Es decir, $\ker T = \{\mathbf{0}\}$ y T es 1-1. Si $p(x) = a_0 + a_1x + a_2x^2$, entonces $p(x) = T\begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix}$. Esto significa que imag $T = P_2$ y T es sobre. De esta manera $\mathbb{R}^3 \cong P_2$.

□ **Ejemplo 5** Sea $V = \{f \in C'[0, 1]: f(0) = 0\}$ y $W = C[0, 1]$. Sea $D: V \to W$ dada por $Df = f'$. Suponga que $Df = Dg$. Entonces $f' = g'$ o $(f - g)' = 0$ y $f(x) - g(x) = c$, una constante. Pero $f(0) = g(0) = 0$, por lo que $c = 0$ y $f = g$. De esta manera D es 1-1. Sea $g \in C[0, 1]$ y sea $f(x) = \int_0^x g(t)\,dt$. Entonces, del Teorema Fundamental del Cálculo, $f \in C'[0, 1]$ y $f'(x) = g(x)$ para toda $x \in [0, 1]$. Más aún, puesto que $\int_0^0 g(t)\,dt = 0$, tenemos $f(0) = 0$. De esta manera, para toda g en W, existe una $f \in V$ tal que $Df = g$. Por esto D es sobre y hemos mostrado que $V \cong W$.

El siguiente teorema muestra la analogía de dos espacios vectoriales isomorfos.

Teorema 5 Sea $T: V \to W$ un isomorfismo.

i. Si $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ genera V, entonces $T\mathbf{v}_1, T\mathbf{v}_2, \ldots, T\mathbf{v}_n$ genera W.
ii. Si $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ son linealmente independientes en V, entonces $T\mathbf{v}_1, T\mathbf{v}_2, \ldots, T\mathbf{v}_n$ son linealmente independientes en W.
iii. Si $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ es una base en V, entonces $\{T\mathbf{v}_1, T\mathbf{v}_2, \ldots, T\mathbf{v}_n\}$ es una base en W.
iv. Si V es de dimensión finita, entonces W es de dimensión finita y $\dim V = \dim W$.

Demostración

i. Sea $\mathbf{w} \in W$. Entonces, puesto que T es sobre, existe una $\mathbf{v} \in V$ tal que $T\mathbf{v} = \mathbf{w}$. Puesto que las $\mathbf{v}_i$ generan V, podemos escribir $\mathbf{v} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_n\mathbf{v}_n$ por lo que $\mathbf{w} = T\mathbf{v} = a_1T\mathbf{v}_1 + a_2T\mathbf{v}_2 + \cdots + a_nT\mathbf{v}_n$ y esto muestra que $\{T\mathbf{v}_1, T\mathbf{v}_2, \ldots, T\mathbf{v}_n\}$ genera W.
ii. Suponiendo que $c_1T\mathbf{v}_1 + c_2T\mathbf{v}_2 + \cdots + c_nT\mathbf{v}_n = \mathbf{0}$. Entonces, si $T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n) = \mathbf{0}$. De esta manera, puesto que T es 1-1, $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n = \mathbf{0}$, lo cual implica que $c_1 = c_2 = \cdots = c_n = 0$ puesto que las $\mathbf{v}_i$ son independientes.
iii. Esto se deduce de las partes (*i*) y (*ii*).
iv. Esto se deduce de la parte (*iii*). ■

En general, es difícil demostrar que dos espacios vectoriales de dimensión infinita son isomorfos. Sin embargo, para espacios de dimensión finita, es muy fácil. El Teorema 3 muestra que si $\dim V \neq \dim W$, entonces V y W no son isomorfos. El siguiente teorema muestra que si $\dim V = \dim W$ y si V y W son espacios vectoriales reales, entonces V y W son isomorfos.

Teorema 6 Sean V y W dos espacios vectoriales de dimensión finita con $\dim V = \dim W$. Entonces $V \cong W$.

Demostración Sea $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ una base para V y sea $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n\}$ una base para W. Defina la transformación lineal T por

$$T\mathbf{v}_i = \mathbf{w}_i \qquad \text{para } i = 1, 2, \ldots, n \tag{2}$$

Por el Teorema 5.2.2 existe exactamente una transformación lineal que satisface la Ecuación (2). Suponga que $\mathbf{v} \in V$ y $T\mathbf{v} = \mathbf{0}$. Entonces, si $\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n$, tenemos $T\mathbf{v} = c_1T\mathbf{v}_1 + \cdots + c_nT\mathbf{v}_n = c_1\mathbf{w}_1 + c_2\mathbf{w}_2 + \cdots + c_n\mathbf{w}_n = \mathbf{0}$. Pero, puesto que $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n$ son linealmente independientes, $c_1 = c_2 = \cdots = c_n = 0$. De esta manera $\mathbf{v} = \mathbf{0}$ y T es 1-1. Puesto que V y W son de dimensión finita y $\dim V = \dim W$, T es sobre por el Teorema 2 y la demostración está completa. ■

El último resultado en esta sección utiliza el material de la Sección 1.11. En esta última definimos las inversas derechas e izquierdas de una matriz rectangular. En el Teorema 5 se vio que la transformación lineal $T: V \to W$ no puede ser un isomorfismo si $\dim V \neq \dim W$. Sin embargo, puede ser una transformación 1-1 o una suprayectiva.

La demostración del siguiente teorema se deja como ejercicio (vea Problemas 21 y 22).

Teorema 7 Sea $T: V \to W$ una transformación lineal con $\dim V = n$ y $\dim W = m$. Sea A_T la representación de la matriz de $m \times n$ de T con respecto a las bases B_1 de V y B_2 de W. Entonces

i. T es suprayectiva si y sólo si A_T tiene inversa derecha.
ii. T es 1-1 si A_T tiene inversa izquierda.

Problemas 5.4

1. Muestre que $T: M_{mn} \to M_{nm}$ definida por $TA = A^t$ es un isomorfismo.
2. Muestre que $T: \mathbb{R}^n \to \mathbb{R}^n$ es un isomorfismo si y sólo si A_T es invertible.
★ 3. Sean V y W espacios vectoriales reales de dimensión n y sean B_1 y B_2 bases para V y W, respectivamente. Sea A_T la matriz transformación relativa a las bases B_1 y B_2. Muestre que $T: V \to W$ es un isomorfismo si y sólo si $\det A_T \neq 0$.
4. Encuentre un isomorfismo entre D_n, la matriz diagonal de $n \times n$ con entradas reales, y $\mathbb{R}^n$. [*Sugerencia:* Observe primero el caso $n = 2$.]
5. ¿Para qué valor de m el conjunto de matrices simétricas de $n \times n$ es isomorfa a $\mathbb{R}^m$?
6. Muestre que el conjunto de matrices simétricas de $n \times n$ es isomorfo al conjunto de matrices triangulares superiores de $n \times n$.
7. Sea $V = P_4$ y $W = \{p \in P_5: p(0) = 0\}$. Muestre que $V \cong W$.
□ 8. Sea $T: P_n \to P_n$ definida por $Tp = p + p'$. Demuestre que T es un isomorfismo.
9. Encuentre una condición para los números m, n, p, q tal que $M_{mn} \cong M_{pq}$.
10. Muestre que $D_n \cong P_{n-1}$.

11. Demuestre que dos espacios vectoriales complejos cualesquiera V y W de dimensión finita con $\dim V = \dim W$ son isomorfos.
12. Sea $T: C[0, 1] \to C[3, 4]$ definida por $Tf(x) = f(x-3)$. Muestre que T es un isomorfismo.
13. Sea B una matriz invertible de $n \times n$. Muestre que $T: M_{nm} \to M_{nm}$ definida por $TA = AB$ es un isomorfismo.
□ 14. Muestre que la transformación $Tp(x) = xp'(x)$ no es un isomorfismo de P_n en P_n.
15. Sea H un subespacio del espacio producto interior V de dimensión finita. Muestre que $T: V \to H$ definida por $T\mathbf{v} = \text{proy}_H \mathbf{v}$ es sobre. ¿En qué condiciones será 1-1?
16. Demuestre que si $T: V \to W$ es un isomorfismo entonces existe un isomorfismo. $S: W \to V$ tal que $S(T\mathbf{v}) = \mathbf{v}$. Aquí S es llamada la *transformación inversa* de T y se denota T^{-1}.
17. Demuestre que si $T: \mathbb{R}^n \to \mathbb{R}^n$ está definida por $T\mathbf{x} = A\mathbf{x}$, y si T es un isomorfismo, entonces A es invertible y la transformación inversa T^{-1} está dada por $T^{-1}\mathbf{x} = A^{-1}\mathbf{x}$.
18. Encuentre T^{-1} para el isomorfismo del Problema 7.
★ 19. Considere el espacio $C = \{z = a + ib$, donde a y b son números reales e $i^2 = -1\}$. Muestre que si los escalares se consideran reales, entonces $C \cong \mathbb{R}^2$.
★ 20. Considere el espacio $\mathbb{C}^n_{\mathbb{R}} = \{(\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n)$: $\mathbf{c}_i \in C$ y los escalares son los reales$\}$. Muestre que $\mathbb{C}^n_{\mathbb{R}} \cong \mathbb{R}^{2n}$. [*Sugerencia:* Vea Problema 19.]
★ 21. Demuestre el Teorema 7(i). [*Sugerencia:* Use el Teorema 1.11.1'.]
★ 22. Demuestre el Teorema 7(ii). [*Sugerencia:* Use el Teorema 1.11.2'.]

5.5 Isometrías (opcional)

En esta sección describiremos un caso especial de transformación lineal entre espacios vectoriales. Comenzaremos con un resultado muy útil.

Teorema 1 Sea A una matriz de $m \times n$ con componentes reales.* Entonces para dos vectores $\mathbf{x} \in \mathbb{R}^n$ y $\mathbf{y} \in \mathbb{R}^m$ cualesquiera:

$$(A\mathbf{x}) \cdot \mathbf{y} = \mathbf{x} \cdot (A^t\mathbf{y}) \qquad (1)$$

Demostración Demostraremos la Ecuación (1) por un método directo, calculando ambos lados separadamente y haciendo ver que son iguales. Así

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \qquad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \qquad \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}$$

* Este resultado puede extenderse fácilmente a matrices con componentes completas. (Vea Problema 21.)

Entonces

$$A\mathbf{x} = \begin{pmatrix} \mathbf{r}_1 \cdot \mathbf{x} \\ \mathbf{r}_2 \cdot \mathbf{x} \\ \vdots \\ \mathbf{r}_m \cdot \mathbf{x} \end{pmatrix}$$

donde $\mathbf{r}_i$ denota el i-ésimo renglón de A y

$$(A\mathbf{x}) \cdot \mathbf{y} = (\mathbf{r}_1 \cdot \mathbf{x})y_1 + (\mathbf{r}_2 \cdot \mathbf{x})y_2 + \cdots + (\mathbf{r}_m \cdot \mathbf{x})y_m = \sum_{i=1}^{m} (\mathbf{r}_i \cdot \mathbf{x})y_i \qquad (2)$$

Pero $\mathbf{r}_i \cdot \mathbf{x} = a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = \sum_{j=1}^{n} a_{ij}x_j$. De esta manera de (2), obtenemos

$$(A\mathbf{x}) \cdot \mathbf{y} = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}x_jy_i \qquad (3)$$

Ahora

$$A^t = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{pmatrix}$$

por lo que

$$A^t\mathbf{y} = \begin{pmatrix} \mathbf{c}_1 \cdot \mathbf{y} \\ \mathbf{c}_2 \cdot \mathbf{y} \\ \vdots \\ \mathbf{c}_n \cdot \mathbf{y} \end{pmatrix}$$

donde $\mathbf{c}_j$ denota la j-ésima columna de A ($=$ j-ésimo renglón de A^t). Entonces

$$(A^t\mathbf{y}) \cdot \mathbf{x} = (\mathbf{c}_1 \cdot \mathbf{y})x_1 + (\mathbf{c}_2 \cdot \mathbf{y})x_2 + \cdots + (\mathbf{c}_n \cdot \mathbf{y})x_n = \sum_{j=1}^{n} (\mathbf{c}_j \cdot \mathbf{y})x_j \qquad (4)$$

Pero $\mathbf{c}_j \cdot \mathbf{y} = a_{1j}y_1 + a_{2j}y_2 + \cdots + a_{mj}y_m = \sum_{i=1}^{m} a_{ij}y_i$. De esta manera, de (4) obtenemos

$$(A^t\mathbf{y}) \cdot \mathbf{x} = \sum_{j=1}^{n} \sum_{i=1}^{m} a_{ij}y_ix_j = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}x_jy_i = (A\mathbf{x}) \cdot \mathbf{y}$$

y el teorema está demostrado. ■

Recuerde de la Sección 4.10 que una matriz Q con componentes reales es *ortogonal* si Q es invertible y $Q^{-1} = Q^t$. En el Teorema 4.10.3 demostramos que Q es ortogonal si y sólo si las columnas de Q forman una base ortonormal de $\mathbb{R}^n$. Ahora, sea Q una matriz ortogonal de $n \times n$ y sea T: $\mathbb{R}^n \to \mathbb{R}^n$ la transformación lineal definida por $T\mathbf{x} = Q\mathbf{x}$. Entonces, usando la Ecuación (1) calculamos $(T\mathbf{x} \cdot T\mathbf{y}) = Q\mathbf{x} \cdot Q\mathbf{y} = \mathbf{x} \cdot (Q^tQ\mathbf{y}) = \mathbf{x} \cdot (I\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$. En particular, si $\mathbf{x} = \mathbf{y}$, vemos que $T\mathbf{x} \cdot T\mathbf{x} = \mathbf{x} \cdot \mathbf{x}$ o bien

$$|T\mathbf{x}| = |\mathbf{x}| \tag{5}$$

para toda $\mathbf{x}$ en $\mathbb{R}^n$.

Definición 1 ***Isometría.*** Una transformación lineal T: $\mathbb{R}^n \to \mathbb{R}^n$ se llama *isometría* si para toda $\mathbf{x}$ y $\mathbf{y}$ en $\mathbb{R}^n$,

$$T\mathbf{x} \cdot T\mathbf{y} = \mathbf{x} \cdot \mathbf{y} \tag{6}$$

Debido a la Ecuación (6) podemos decir: *Una isometría en $\mathbb{R}^n$ es una transformación lineal que conserva el producto escalar en $\mathbb{R}^n$.*

Ejemplo 1 Sea T: $\mathbb{R}^2 \to \mathbb{R}^2$ un operador lineal que hace girar a cada vector en $\mathbb{R}^2$ un ángulo θ. Entonces, como vimos en el Ejemplo 5.1.8, $T\mathbf{x} = A_\theta\mathbf{x}$, donde $A_\theta = \begin{pmatrix} \cos\theta & -\text{sen}\,\theta \\ \text{sen}\,\theta & \cos\theta \end{pmatrix}$. Pero A_θ es una matriz ortogonal, por lo que T es una isometría. Para verificar note que $T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x\cos\theta - y\,\text{sen}\,\theta \\ x\,\text{sen}\,\theta + y\cos\theta \end{pmatrix}$ y $\left|T\begin{pmatrix} x \\ y \end{pmatrix}\right|^2 =$ $\left[T\begin{pmatrix} x \\ y \end{pmatrix}\right] \cdot \left[T\begin{pmatrix} x \\ y \end{pmatrix}\right] = [x\cos\theta - y\,\text{sen}\,\theta]^2 + [x\,\text{sen}\,\theta + y\cos\theta]^2 =$
$(x^2\cos^2\theta - 2xy\cos\theta\,\text{sen}\,\theta + y^2\,\text{sen}^2\,\theta) +$
$(x^2\,\text{sen}^2\,\theta + 2xy\,\text{sen}\,\theta\cos\theta + y^2\cos^2\theta) =$
$x^2(\cos^2\theta + \text{sen}^2\,\theta) + y^2(\text{sen}^2\,\theta + \cos^2\theta) = x^2 + y^2 = \left|\begin{pmatrix} x \\ y \end{pmatrix}\right|^2.$
De esta manera $\left|T\begin{pmatrix} x \\ y \end{pmatrix}\right| = \left|\begin{pmatrix} x \\ y \end{pmatrix}\right|.$

Las isometrías tienen algunas propiedades interesantes.

Teorema 2 Sea T: $\mathbb{R}^n \to \mathbb{R}^n$ una isometría. Entonces

i. Si $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ es un conjunto ortogonal entonces $T\mathbf{u}_1, T\mathbf{u}_2, \ldots, T\mathbf{u}_n$ es un conjunto ortogonal.
ii. T es un isomorfismo.

Demostración **i.** Si $i \neq j$ y $\mathbf{u}_i \cdot \mathbf{u}_j = 0$, entonces $(T\mathbf{u}_i) \cdot (T\mathbf{u}_j) = \mathbf{u}_i \cdot \mathbf{u}_j = 0$, lo cual prueba (i).
ii. Sea $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ una base ortonormal para $\mathbb{R}^n$. Entonces por la parte (i) y el hecho de que $|T\mathbf{u}_i| = |\mathbf{u}_i| = 1$, tenemos que $T\mathbf{u}_1, T\mathbf{u}_2, \ldots, T\mathbf{u}_n$ es un conjunto ortonormal en $\mathbb{R}^n$. Por el Teorema 4.10.1 estos vectores son linealmente independientes y por lo tanto forman una base de $\mathbb{R}^n$. De esta manera imag $T = \mathbb{R}^n$, lo cual prueba que ker $T = \{\mathbf{0}\}$ (puesto que $\nu(T) + \rho(T) = n$). ■

Teorema 3 Sea $B = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ una base ortonormal de $\mathbb{R}^n$. Entonces T: $\mathbb{R}^n \to \mathbb{R}^n$ es una isometría si y sólo si A_T es ortogonal, donde A_T es la representación matricial de T relativa a la base B.

Demostración Suponga que A_T es ortogonal con respecto a B. Entonces, como hemos visto con anterioridad, $T\mathbf{x} \cdot T\mathbf{x} = A_T\mathbf{x} \cdot A_T\mathbf{x} = \mathbf{x} \cdot \mathbf{x}$, por lo que T es una isometría. Si T es una isometría, entonces $\mathbf{x} \cdot \mathbf{y} = T\mathbf{x} \cdot T\mathbf{y} = A_T\mathbf{x} \cdot A_T\mathbf{y} = \mathbf{x} \cdot A_T^tA_T\mathbf{y}$. Esto significa que para todos los vectores $\mathbf{x}$ y $\mathbf{y}$ en $\mathbb{R}^n$, $\mathbf{x} \cdot (\mathbf{y} - A_T^tA_T\mathbf{y}) = 0$. Por lo tanto $(\mathbf{y} - A_T^tA_T) \in (\mathbb{R}^n)^\perp = \{\mathbf{0}\}$. De esta manera $\mathbf{y} = A_T^tA_T\mathbf{y}$ para todo $\mathbf{y} \in \mathbb{R}^n$ lo cual significa que $A_T^tA_T = I$ y $A_T^t = A_T^{-1}$. De esta manera A_T es ortogonal y el teorema está demostrado. ■

Concluimos esta sección indicando cómo podemos extender el concepto de isometría a un espacio producto interior arbitrario.

Definición 2 ***Isometría.*** Sean V y W espacios producto interior reales (o complejos) y sea T: $V \to W$ una transformación lineal. Entonces T es una *isometría* si para toda $\mathbf{v}_1$, $\mathbf{v}_2 \in V$,

$$(\mathbf{v}_1, \mathbf{v}_2) = (T\mathbf{v}_1, T\mathbf{v}_2) \tag{7}$$

Nota. Por la Ecuación (7), si T es una isometría entonces (usando la Ecuación 4.11.3) vemos que

$$|\mathbf{v}_1|_V = |T\mathbf{v}_1|_W \tag{8}$$

Definición 3 ***Espacios vectoriales isométricamente isomórficos.*** Se dice que dos espacios vectoriales V y W son *isométricamente isomorfos* si existe una transformación lineal T: $V \to W$ que sea una isometría y un isomorfismo al mismo tiempo.

Teorema 4 Dos espacios producto interior reales cualesquiera de dimensión n son isométricamente isomorfos.

Demostración Sean $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ y $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n\}$ bases ortonormales de V y W, respectivamente. Sea $T: V \to W$ la transformación lineal definida por $T\mathbf{u}_i = \mathbf{w}_i$, $i = 1, 2, \ldots, n$. Si podemos demostrar que T es una isometría, entonces habremos terminado pues, razonando como en la demostración del Teorema 2, podemos ver que T también es un isomorfismo. Sean $\mathbf{x}$ y $\mathbf{y}$ en V. Entonces existen conjuntos de números reales $c_1, c_2, \ldots, c_n$ y $d_1, d_2, \ldots, d_n$ tales que $\mathbf{x} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_n\mathbf{u}_n$ y $\mathbf{y} = d_1\mathbf{u}_1 + d_2\mathbf{u}_2 + \cdots + d_n\mathbf{u}_n$.

Puesto que las $\mathbf{u}_i$ son ortonormales, $(\mathbf{x}, \mathbf{y}) = [(c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_n\mathbf{u}_n), (d_1\mathbf{u}_1 + d_2\mathbf{u}_2 + \cdots + d_n\mathbf{u}_n)] = c_1d_1 + c_2d_2 + \cdots + c_nd_n$. Análogamente puesto que $T\mathbf{x} = c_1T\mathbf{u}_1 + c_2T\mathbf{u}_2 + \cdots + c_nT\mathbf{u}_n = c_1\mathbf{w}_1 + c_2\mathbf{w}_2 + \cdots + c_n\mathbf{w}_n$, entonces obtenemos $(T\mathbf{x}, T\mathbf{y}) = [(c_1\mathbf{w}_1 + c_2\mathbf{w}_2 + \cdots + c_n\mathbf{w}_n), (d_1\mathbf{w}_1 + d_2\mathbf{w}_2 + \cdots + d_n\mathbf{w}_n)] = c_1d_1 + c_2d_2 + \cdots + c_nd_n$ pues las $\mathbf{w}_i$ son ortonormales. Esto completa la demostración. ■

□ ***Ejemplo 2*** Ilustramos este teorema mostrando que $\mathbb{R}^3$ y $P_2[0, 1]$ son isométricamente isomorfos. En $\mathbb{R}^3$ usamos la base estándar $\left\{\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix}\right\}$. En P_2 usamos la base ortonormal $\{1, \sqrt{3}(2x-1), \sqrt{5}(6x^2-6x+1)\}$. (Vea Ejemplo 4.11.8.) Sean $\mathbf{x} = \begin{pmatrix}a_1\\b_1\\c_1\end{pmatrix}$ y $\mathbf{y} = \begin{pmatrix}a_2\\b_2\\c_2\end{pmatrix}$ en $\mathbb{R}^3$. Entonces $(\mathbf{x}, \mathbf{y}) = \mathbf{x}\cdot\mathbf{y} = a_1a_2 + b_1b_2 + c_1c_2$. Recuerde que en $P_2[0, 1]$ definimos $(p, q) = \int_0^1 p(x)q(x)\,dx$. Ahora $T\begin{pmatrix}1\\0\\0\end{pmatrix} = 1$,

$T\begin{pmatrix}0\\1\\0\end{pmatrix} = \sqrt{3}(2x-1)$, y $T\begin{pmatrix}0\\0\\1\end{pmatrix} = \sqrt{5}(6x^2-6x+1)$; por lo tanto $T\begin{pmatrix}a\\b\\c\end{pmatrix} = a + b\sqrt{3}(2x-1) + c\sqrt{5}(6x^2-6x+1)$ y

$$(T\mathbf{x}, T\mathbf{y}) = \int_0^1 [a_1 + b_1\sqrt{3}(2x-1) + c_1\sqrt{5}(6x^2-6x+1)]$$

$$\times[a_2 + b_2\sqrt{3}(2x-1) + c_2\sqrt{5}(6x^2-6x+1)]\,dx$$

$$= a_1a_2\int_0^1 dx + \int_0^1 b_1b_2 3(2x-1)^2\,dx + \int_0^1 c_1c_2[5(6x^2-6x+1)^2]\,dx$$

$$+(a_1b_2 + a_2b_1)\int_0^1 \sqrt{3}(2x-1)\,dx$$

$$+(a_1c_2 + a_2c_1)\int_0^1 \sqrt{5}(6x^2-6x+1)\,dx$$

$$+(b_1c_2 + b_2c_1)\int_0^1 [\sqrt{3}(2x-1)][\sqrt{5}(6x^2-6x+1)]\,dx$$

$$= a_1a_2 + b_1b_2 + c_1c_2$$

Aquí hemos ahorrado tiempo pues usamos el hecho de que $\{1, \sqrt{3}(2x-1), \sqrt{5}(6x^2-6x+1)\}$ es una base ortonormal. De esta manera $T: \mathbb{R}^3 \to P_2[0, 1]$ es una isometría.

Problemas 5.5

1. Muestre que para todo número real θ, la transformación $T: \mathbb{R}^3 \to \mathbb{R}^3$ definida por $T\mathbf{x} = A\mathbf{x}$, donde

$$A = \begin{pmatrix} \operatorname{sen}\theta & \cos\theta & 0 \\ \cos\theta & -\operatorname{sen}\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

es una isometría.

2. Haga lo mismo para la transformación T donde

$$A = \begin{pmatrix} \cos\theta & 0 & -\operatorname{sen}\theta \\ 0 & 1 & 0 \\ \operatorname{sen}\theta & 0 & \cos\theta \end{pmatrix}$$

3. Sean A y B matrices ortogonales de $n \times n$. Muestre que $T: \mathbb{R}^n \to \mathbb{R}^n$ definida por $T\mathbf{x} = AB\mathbf{x}$ es una isometría.
4. Encuentre A_T si T es la transformación de $\mathbb{R}^3 \to \mathbb{R}^3$ definida por

$$T\begin{pmatrix}2/3\\1/3\\-2/3\end{pmatrix} = \begin{pmatrix}1/\sqrt{2}\\1/\sqrt{2}\\0\end{pmatrix} \quad T\begin{pmatrix}1/3\\2/3\\2/3\end{pmatrix} = \begin{pmatrix}-1/\sqrt{6}\\1/\sqrt{6}\\2/\sqrt{6}\end{pmatrix} \quad T\begin{pmatrix}2/3\\-2/3\\1/3\end{pmatrix} = \begin{pmatrix}1/\sqrt{3}\\-1/\sqrt{3}\\1/\sqrt{3}\end{pmatrix}$$

Demuestre que A_T es ortogonal.

5. Sea $T: \mathbb{R}^n \to \mathbb{R}^n$ con la propiedad de que $|T\mathbf{x}| = |\mathbf{x}|$ para toda $\mathbf{x} \in \mathbb{R}^n$. Demuestre que T es una isometría. [*Sugerencia:* Muestre primero que $\mathbf{x}\cdot\mathbf{y} = \frac{1}{2}(|\mathbf{x}+\mathbf{y}|^2 - |\mathbf{x}|^2 - |\mathbf{y}|^2)$.]
6. Sea $T: \mathbb{R}^2 \to \mathbb{R}^2$ una isometría. Muestre que T conserva los ángulos. Esto es: (ángulo entre $\mathbf{x}$ y $\mathbf{y}$) = (ángulo entre $T\mathbf{x}$ y $T\mathbf{y}$).
7. Dé un ejemplo de una transformación lineal de $\mathbb{R}^2$ a $\mathbb{R}^2$ que conserve los ángulos y que *no* sea una isometría.
8. Para $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ y $\mathbf{x}$ y $\mathbf{y} \neq \mathbf{0}$, defina: (ángulo entre $\mathbf{x}$ y $\mathbf{y}$) $= \measuredangle(\mathbf{x}, \mathbf{y}) = \cos^{-1}[(\mathbf{x}\cdot\mathbf{y})/|\mathbf{x}|\,|\mathbf{y}|]$. Muestre que si $T: \mathbb{R}^n \to \mathbb{R}^n$ es una isometría, entonces T conserva los ángulos.
9. Sea $T: \mathbb{R}^n \to \mathbb{R}^n$ una isometría y sea $T\mathbf{x} = A\mathbf{x}$. Muestre que $S\mathbf{x} = A^{-1}\mathbf{x}$ es una isometría.

En los Problemas del 10 *al* 14 *encuentre una isometría entre los pares de espacios dados.*

□**10.** $P_1[-1, 1], \mathbb{R}^2$ □★**11.** $P_3[-1, 1], \mathbb{R}^4$ ★**12.** $M_{22}, \mathbb{R}^4$ □★**13.** $M_{22}, P_3[-1, 1]$

14. D_n y $\mathbb{R}^n$ (D_n = conjunto de matrices diagonales de $n \times n$).
15. Sea A una matriz de $n \times n$ con componentes complejas. Entonces el conjugado de A, denotado A^*, se define como $(A^*)_{ij} = \overline{a_{ji}}$. Calcule A^* si $A = \begin{pmatrix}1+i & -4+2i\\ 3 & 6-3i\end{pmatrix}$.

16. La matriz compleja A de $n \times n$ se llama *hermitiana* si $A^* = A$. Muestre que la matriz $A = \begin{pmatrix} 4 & 3-2i \\ 3+2i & 6 \end{pmatrix}$ es hermitiana.

17. Muestre que si A es hermitiana, entonces las componentes diagonales de A son reales.

18. La matriz compleja A de $n \times n$ se llama *unitaria* si $A^* = A^{-1}$. Muestre que la matriz

$$A = \begin{pmatrix} \dfrac{1+i}{2} & \dfrac{3-2i}{\sqrt{26}} \\ \dfrac{1+i}{2} & \dfrac{-3+2i}{\sqrt{26}} \end{pmatrix}$$

es unitaria.

19. Muestre que A es unitaria si y sólo si las columnas de A forman una base ortonormal en $\mathbb{C}^n$.

20. Muestre que si A es unitaria entonces $|\det A| = 1$.

21. Sea A una matriz de $n \times n$ con componentes complejas. En $\mathbb{C}^n$, si $\mathbf{x} = (c_1, c_2, \ldots, c_n)$ y $\underline{y} = (d_1, d_2, \ldots, d_n)$, defina el producto interior $(\mathbf{x}, \mathbf{y}) = c_1\bar{d}_1 + c_2\bar{d}_2 + \ldots + c_n\bar{d}_n$. (Ejemplo 4.11.2.) Demuestre que $(A\mathbf{x}, \mathbf{y}) = (\mathbf{x}, A^*\mathbf{y})$.

★ 22. Muestre que dos espacios producto interior complejos cualesquiera de la misma dimensión (finita) son isométricamente isomorfos.

Ejercicios de repaso • Capítulo 5

En los Ejercicios del 1 al 6 determine si la transformación dada de V a W es lineal.

1. T: $\mathbb{R}^2 \to \mathbb{R}^2$; $T(x, y) = (0, -y)$
2. T: $\mathbb{R}^3 \to \mathbb{R}^3$; $T(x, y, z) = (1, y, z)$
3. T: $\mathbb{R}^2 \to \mathbb{R}^2$; $T(x, y) = x/y$
4. T: $P_1 \to P_2$; $(Tp)(x) = xp(x)$
5. T: $P_2 \to P_2$; $(Tp)(x) = 1 + p(x)$
6. T: $C[0, 1] \to C[0, 1]$; $Tf(x) = f(1)$

En los Ejercicios del 7 al 12 halle la representación matricial de la transformación lineal dada y encuentre el núcleo o kernel, la imagen, la nulidad y el rango de la transformación.

7. T: $\mathbb{R}^2 \to \mathbb{R}^2$; $T(x, y) = (0, -y)$
8. T: $\mathbb{R}^3 \to \mathbb{R}^2$; $T(x, y, z) = (y, z)$
9. T: $\mathbb{R}^4 \to \mathbb{R}^2$; $T(x, y, z, w) = (x - 2z, 2y + 3w)$
10. T: $P_3 \to P_4$; $(Tp)(x) = xp(x)$
11. T: $M_{22} \to M_{22}$; $TA = AB$, en donde $B = \begin{pmatrix} -1 & 0 \\ 1 & 2 \end{pmatrix}$
12. T: $\mathbb{R}^2 \to \mathbb{R}^2$; $T(x, y) = (x - y, 2x + 3y)$; $B_1 = \left\{\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix}\right\}$; $B_2 = \left\{\begin{pmatrix} -1 \\ 3 \end{pmatrix}, \begin{pmatrix} 4 \\ 1 \end{pmatrix}\right\}$
13. Encuentre el isomorfismo de $T: P_2 \to \mathbb{R}^3$.
14. Encuentre la isometría de $T: \mathbb{R}^2 \to P_1[-1, 1]$.

CAPÍTULO **6**

Valores característicos, vectores característicos y formas canónicas

6.1 Valores característicos y vectores característicos

Sea $T: V \to V$ una transformación lineal. En una gran variedad de aplicaciones (una de las cuales se da en la siguiente sección), resulta útil encontrar un vector $\mathbf{v}$ en V tal que $T\mathbf{v}$ y $\mathbf{v}$ sean paralelos. Esto es, buscamos un vector $\mathbf{v}$ y un escalar λ tal que

$$T\mathbf{v} = \lambda\mathbf{v} \tag{1}$$

Si $\mathbf{v} \neq \mathbf{0}$ y λ satisface (1) entonces λ se denomina *valor característico* de T y $\mathbf{v}$ se llama *vector característico* de T correspondiente al valor característico λ. El propósito de este capítulo es el de investigar algunas propiedades de los valores y vectores característicos. Si V es de dimensión finita, entonces T puede representarse por una matriz A_T. Por esta razón discutiremos valores y vectores característicos de matrices de $n \times n$.

Definición 1 **Valor característico y vector característico.** Sea A una matriz de $n \times n$ con componentes reales*. El número λ (real o complejo) se llama *valor característico* de A si hay un vector $\mathbf{v}$ *distinto de cero* en $\mathbb{C}^n$ tal que

$$A\mathbf{v} = \lambda\mathbf{v} \tag{2}$$

* Esta definición también es válida si A tiene componentes complejas; pero como las matrices que utilizaremos en la mayoría de los casos tienen componentes reales, la definición será suficiente para nuestros propósitos.

El vector $\mathbf{v} \neq \mathbf{0}$ se llama un *vector característico de A correspondiente al valor característico* λ.

Nota. La palabra *eigen* significa "propio" o "apropiado" en alemán. Los valores característicos se llaman también *valores propios* o *autovalores*, y los vectores característicos, *vectores propios* o *autovectores*.

Observación. Como veremos (en el Ejemplo 6) una matriz con componentes reales puede tener valores y vectores característicos complejos. Esta es la razón por la que dijimos que $\mathbf{v} \in \mathbb{C}^n$. En este libro no se utilizarán muchos resultados acerca de los números complejos. Para una discusión de los resultados que se utilizarán, vea el Apéndice 2.

Ejemplo 1 Sea $A = \begin{pmatrix} 10 & -18 \\ 6 & -11 \end{pmatrix}$. Entonces $A\begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 10 & -18 \\ 6 & -11 \end{pmatrix}\begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$. De esta manera $\lambda_1 = 1$ es un valor característico de A correspondiente al vector característico $\mathbf{v}_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$. Análogamente, $A\begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 10 & -18 \\ 6 & -11 \end{pmatrix}\begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} -6 \\ -4 \end{pmatrix} = -2\begin{pmatrix} 3 \\ 2 \end{pmatrix}$, por lo que $\lambda_2 = -2$ es un valor característico de A con su correspondiente vector característico $\mathbf{v}_2 = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$. Como pronto veremos, estos son los únicos valores característicos de A.

Ejemplo 2 Sea $A = I$. Entonces para todo $\mathbf{v} \in \mathbb{C}^n$, $A\mathbf{v} = I\mathbf{v} = \mathbf{v}$. De esta manera 1 es el único valor característico de A y cada $\mathbf{v} \neq \mathbf{0} \in \mathbb{C}^n$ es un vector característico de I.

En esta sección calcularemos los valores y vectores característicos de muchas matrices. Pero primero necesitamos demostrar algunos hechos que simplificarán nuestros cálculos.

Suponga que λ es un valor característico de A. Entonces existe un vector no nulo $\mathbf{v} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \neq \mathbf{0}$ tal que $A\mathbf{v} = \lambda\mathbf{v} = \lambda I\mathbf{v}$. Volviendo a escribir esto tenemos

$$(A - \lambda I)\mathbf{v} = \mathbf{0} \tag{3}$$

Si A es una matriz de $n \times n$, la Ecuación (3) es un sistema homogéneo de n ecuaciones en las incógnitas $x_1, x_2, \ldots, x_n$. Puesto que, por hipótesis, el sistema tiene soluciones no triviales, concluimos que $\det(A - \lambda I) = 0$. Inversamente, si $\det(A - \lambda I) = 0$, entonces la Ecuación (3) tiene soluciones no triviales y λ es un valor característico de A. Por otro lado, si $\det(A - \lambda I) \neq 0$, entonces (3) tiene únicamente la solución $\mathbf{v} = \mathbf{0}$ por lo que λ *no* es un valor característico de A. Resumiendo estos resultados tenemos el siguiente teorema.

Teorema 1 Sea A una matriz de $n \times n$. Entonces λ es un valor característico de A si y sólo si

$$p(\lambda) = \det(A - \lambda I) = 0 \tag{4}$$

Definición 2 **Ecuación característica y polinomio característico.** La Ecuación (4) se conoce como la *ecuación característica* de A; $p(\lambda)$ se conoce como el *polinomio característico* de A.

De los ejemplos resultará evidente que $p(\lambda)$ es un polinomio de grado n en λ. Por ejemplo, si $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, entonces $A - \lambda I = \begin{pmatrix} a & b \\ c & d \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} = \begin{pmatrix} a-\lambda & b \\ c & d-\lambda \end{pmatrix}$ y $p(\lambda) = \det(A - \lambda I) = (a - \lambda)(d - \lambda) - bc = \lambda^2 - (a + d)\lambda + (ad - bc)$.

Por el *teorema fundamental del álgebra*, todo polinomio de grado n con coeficientes reales o complejos tiene n raíces exactamente (contando multiplicidades). Con esto queremos decir que, por ejemplo, el polinomio $(\lambda - 1)^5$ tiene cinco raíces, todas iguales al número 1. Puesto que todo valor característico de A es una raíz de la ecuación característica de A, concluimos que:

Si se consideran multiplicidades, cada matriz de $n \times n$ tiene exactamente n valores característicos.

Teorema 2 Sea λ un valor característico de la matriz A de $n \times n$ y sea $E_\lambda = \{\mathbf{v}: A\mathbf{v} = \lambda\mathbf{v}\}$. Entonces E_λ es un subespacio de $\mathbb{C}^n$.

Demostración Si $A\mathbf{v} = \lambda\mathbf{v}$, entonces $(A - \lambda I)\mathbf{v} = \mathbf{0}$. De esta manera E_λ es el núcleo de la matriz $A - \lambda I$, la cual, por el Ejemplo 4.6.10, es un subespacio* de $\mathbb{C}^n$. ■

Definición 3 **Espacio característico.** Sea λ un valor característico de A. El subespacio E_λ se denomina *espacio característico*† de A correspondiente al valor característico λ.

Ahora demostraremos otro resultado de gran utilidad.

Teorema 3 Sea A una matriz de $n \times n$ y sean $\lambda_1, \lambda_2, \ldots, \lambda_m$ valores característicos diferentes de A con sus correspondientes vectores característicos $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$. Entonces $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ son linealmente independientes. Esto es: *los vectores característicos correspondientes a valores característicos diferentes son linealmente independientes.*

* En el Ejemplo 4.6.10 vimos que ker A es un subespacio de $\mathbb{R}^n$ si A es una matriz real. La extensión de este resultado a $\mathbb{C}^n$ no presenta dificultades.

† Note que $\mathbf{0} \in E_\lambda$ puesto que E_λ es un subespacio.

Demostración Demostraremos esto por inducción matemática. Comencemos con $m=2$. Suponga que

$$c_1\mathbf{v}_1+c_2\mathbf{v}_2=\mathbf{0} \tag{5}$$

De tal manera, multiplicando ambos lados de (5) por A, tenemos como resultado $\mathbf{0} = A(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) = c_1A\mathbf{v}_1 + c_2A\mathbf{v}_2$ o bien (ya que $A\mathbf{v}_i = \lambda_i\mathbf{v}_i$ para $i = 1, 2$)

$$c_1\lambda_1\mathbf{v}_1+c_2\lambda_2\mathbf{v}_2=\mathbf{0} \tag{6}$$

Luego multiplicamos (5) por λ_1 y restamos esto de (6) para obtener

$$(c_1\lambda_1\mathbf{v}_1+c_2\lambda_2\mathbf{v}_2)-(c_1\lambda_1\mathbf{v}_1+c_2\lambda_1\mathbf{v}_2)=\mathbf{0}$$

o sea

$$c_2(\lambda_2-\lambda_1)\mathbf{v}_2=\mathbf{0}$$

Puesto que $\mathbf{v}_2\neq\mathbf{0}$ (por la definición de vector característico) y ya que $\lambda_2\neq\lambda_1$, concluimos que $c_2=0$. Entonces sustituimos $c_2=0$ en (5) y vemos que $c_1=0$, lo cual demuestra el teorema en el caso $m=2$. Ahora supongamos que el teorema es cierto para $m = k$. Esto es, supongamos que cualesquiera k vectores característicos correspondientes a valores característicos distintos son linealmente independientes. Demostraremos el teorema para $m = k + 1$. Entonces

$$c_1\mathbf{v}_1+c_2\mathbf{v}_2+\cdots+c_k\mathbf{v}_k+c_{k+1}\mathbf{v}_{k+1}=\mathbf{0} \tag{7}$$

De donde, multiplicando ambos lados de (7) por A y usando el hecho de que $A\mathbf{v}_i = \lambda_i\mathbf{v}_i$, resulta

$$c_1\lambda_1\mathbf{v}_1+c_2\lambda_2\mathbf{v}_2+\cdots+c_k\lambda_k\mathbf{v}_k+c_{k+1}\lambda_{k+1}\mathbf{v}_{k+1}=\mathbf{0} \tag{8}$$

Multiplicamos ambos lados de (7) por λ_{k+1} y lo restamos de (8):

$$c_1(\lambda_1-\lambda_{k+1})\mathbf{v}_1+c_2(\lambda_2-\lambda_{k+1})\mathbf{v}_2+\cdots+c_k(\lambda_k-\lambda_{k+1})\mathbf{v}_k=\mathbf{0}$$

Pero, por la hipótesis de inducción $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ son linealmente independientes. De modo que $c_1(\lambda_1 - \lambda_{k+1}) = c_2(\lambda_2 - \lambda_{k+1}) = \cdots = c_k(\lambda_k - \lambda_{k+1}) = 0$; y, puesto que $\lambda_i \neq \lambda_{k+1}$ para $i = 1, 2, \ldots, k$, concluimos que $c_1 = c_2 = \cdots = c_k = 0$. Pero de (7), esto significa que $c_{k+1} = 0$. De esta manera el teorema es cierto para $m = k + 1$ y la demostración es completa. ■

Si

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & & \vdots\\ a_{n1} & a_{n2} & \cdots & a_{nn}\end{pmatrix}$$

entonces

$$p(\lambda)=\det(A-\lambda I)=\begin{vmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22}-\lambda & \cdots & a_{2n}\\ \vdots & \vdots & & \vdots\\ a_{n1} & a_{n2} & \cdots & a_{nn}-\lambda\end{vmatrix}$$

y $p(\lambda) = 0$ puede escribirse en la forma

$$p(\lambda)=\lambda^n+b_{n-1}\lambda^{n-1}+\cdots+b_1\lambda+b_0=0 \tag{9}$$

La Ecuación (9) tiene n raíces, varias de las cuales pueden repetirse. Si $\lambda_1, \lambda_2, \ldots, \lambda_m$ son las diferentes raíces de (9) con multiplicidades $r_1, r_2, \ldots, r_m$, respectivamente, entonces (9) puede factorizarse para obtener

$$p(\lambda)=(\lambda-\lambda_1)^{r_1}(\lambda-\lambda_2)^{r_2}\cdots(\lambda-\lambda_m)^{r_m}=0 \tag{10}$$

Multiplicidad algebraica Los números $r_1, r_2, \ldots, r_m$ se llaman *multiplicidades algebraicas* de los valores característicos $\lambda_1, \lambda_2, \ldots, \lambda_m$, respectivamente.

Procedamos ahora a calcular valores característicos y sus correspondientes espacios característicos. Haremos esto en tres pasos:

Procedimiento para calcular valores propios y vectores propios.

i. Halle $p(\lambda) = \det(A - \lambda I)$.
ii. Halle las raíces $\lambda_1, \lambda_2, \ldots, \lambda_m$ de $p(\lambda) = 0$.
iii. Resuelva el sistema homogéneo $(A - \lambda_i I)\mathbf{v} = \mathbf{0}$ correspondiente a cada valor característico λ_i.

Observación. El paso (*ii*) es, generalmente, el más difícil.

Ejemplo 3 Sea $A = \begin{pmatrix} 4 & 2\\ 3 & 3\end{pmatrix}$. Entonces $\det(A-\lambda I)=\begin{vmatrix} 4-\lambda & 2\\ 3 & 3-\lambda\end{vmatrix}=(4-\lambda)(3-\lambda)-6=\lambda^2-7\lambda+6=(\lambda-1)(\lambda-6)=0$. De esta manera los valores característicos de A son $\lambda_1 = 1$ y $\lambda_2 = 6$. Para $\lambda_1 = 1$, resolvemos $(A - I)\mathbf{v} = 0$, o bien, $\begin{pmatrix} 3 & 2\\ 3 & 2\end{pmatrix}\begin{pmatrix} x_1\\ x_2\end{pmatrix}=\begin{pmatrix} 0\\ 0\end{pmatrix}$. Evidentemente, todo vector característico correspondiente a $\lambda_1=1$ satisface $3x_1+2x_2=0$. Uno de estos vectores característicos es $\mathbf{v}_1=\begin{pmatrix} 2\\ -3\end{pmatrix}$. De esta manera, $E_1 = \text{gen}\left\{\begin{pmatrix} 2\\ -3\end{pmatrix}\right\}$. Análogamente, la ecuación $(A - 6I)\mathbf{v} = \mathbf{0}$ significa que $\begin{pmatrix} -2 & 2\\ 3 & -3\end{pmatrix}\begin{pmatrix} x_1\\ x_2\end{pmatrix}=\begin{pmatrix} 0\\ 0\end{pmatrix}$ o $x_1 = x_2$. Por consiguiente, $\mathbf{v}_2=\begin{pmatrix} 1\\ 1\end{pmatrix}$ es un vector característico correspondiente a $\lambda_2=6$ y $E_6=\text{gen}\left\{\begin{pmatrix} 1\\ 1\end{pmatrix}\right\}$. Note que $\mathbf{v}_1$ y $\mathbf{v}_2$ son linealmente independientes puesto que uno no es múltiplo del otro.

Ejemplo 4 Sea $A = \begin{pmatrix} 1 & -1 & 4 \\ 3 & 2 & -1 \\ 2 & 1 & -1 \end{pmatrix}$. Entonces

$$\det(A-\lambda I) = \begin{vmatrix} 1-\lambda & -1 & 4 \\ 3 & 2-\lambda & -1 \\ 2 & 1 & -1-\lambda \end{vmatrix}$$

$$= -(\lambda^3 - 2\lambda^2 - 5\lambda + 6) = -(\lambda-1)(\lambda+2)(\lambda-3) = 0$$

De esta manera los valores característicos de A son $\lambda_1 = 1$, $\lambda_2 = -2$ y $\lambda_3 = 3$. Para $\lambda_1 = 1$ tenemos

$$(A-I)\mathbf{v} = \begin{pmatrix} 0 & -1 & 4 \\ 3 & 1 & -1 \\ 2 & 1 & -2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Si resolvemos reduciendo por renglones, obtenemos, sucesivamente,

$$\left(\begin{array}{ccc|c} 0 & -1 & 4 & 0 \\ 3 & 1 & -1 & 0 \\ 2 & 1 & -2 & 0 \end{array}\right) \xrightarrow[A_{1,3}(1)]{A_{1,2}(1)} \left(\begin{array}{ccc|c} 0 & -1 & 4 & 0 \\ 3 & 0 & 3 & 0 \\ 2 & 0 & 2 & 0 \end{array}\right) \xrightarrow{M_2(\frac{1}{3})}$$

$$\left(\begin{array}{ccc|c} 0 & -1 & 4 & 0 \\ 1 & 0 & 1 & 0 \\ 2 & 0 & 2 & 0 \end{array}\right) \xrightarrow{A_{2,3}(-2)} \left(\begin{array}{ccc|c} 0 & -1 & 4 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

De esta manera $x_1 = -x_3$, $x_2 = 4x_3$, un vector característico es $\mathbf{v}_1 = \begin{pmatrix} -1 \\ 4 \\ 1 \end{pmatrix}$, y $E_1 = \text{gen}\left\{\begin{pmatrix} -1 \\ 4 \\ 1 \end{pmatrix}\right\}$. Para $\lambda_2 = -2$, tenemos $[A-(-2I)]\mathbf{v} = (A+2I)\mathbf{v} = \mathbf{0}$ o

$$\begin{pmatrix} 3 & -1 & 4 \\ 3 & 4 & -1 \\ 2 & 1 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$ Esto conduce a

$$\left(\begin{array}{ccc|c} 3 & -1 & 4 & 0 \\ 3 & 4 & -1 & 0 \\ 2 & 1 & 1 & 0 \end{array}\right) \xrightarrow[A_{1,3}(1)]{A_{1,2}(4)} \left(\begin{array}{ccc|c} 3 & -1 & 4 & 0 \\ 15 & 0 & 15 & 0 \\ 5 & 0 & 5 & 0 \end{array}\right) \xrightarrow{M_2(\frac{1}{15})}$$

$$\left(\begin{array}{ccc|c} 3 & -1 & 4 & 0 \\ 1 & 0 & 1 & 0 \\ 5 & 0 & 5 & 0 \end{array}\right) \xrightarrow{A_{2,3}(-5)} \left(\begin{array}{ccc|c} -1 & -1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

De esta manera $x_2 = -x_1$, $x_3 = -x_1$, y un vector característico es $\mathbf{v}_2 = \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}$.

Por consiguiente, $E_{-2} = \text{gen}\left\{\begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}\right\}$. Finalmente, para $\lambda_3 = 3$,

$$(A-3I)\mathbf{v} = \begin{pmatrix} -2 & -1 & 4 \\ 3 & -1 & -1 \\ 2 & 1 & -4 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

y

$$\left(\begin{array}{ccc|c} -2 & -1 & 4 & 0 \\ 3 & -1 & -1 & 0 \\ 2 & 1 & -4 & 0 \end{array}\right) \xrightarrow[A_{1,3}(1)]{A_{1,2}(-1)} \left(\begin{array}{ccc|c} -2 & -1 & 4 & 0 \\ 5 & 0 & -5 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

$$\xrightarrow{M_2(\frac{1}{5})} \left(\begin{array}{ccc|c} -2 & -1 & 4 & 0 \\ 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right) \xrightarrow{A_{2,1}(4)} \left(\begin{array}{ccc|c} 2 & -1 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Por lo tanto $x_3 = x_1$, $x_2 = 2x_1$, y $\mathbf{v}_3 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$ por lo que $E_3 = \text{gen}\left\{\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}\right\}$.

Observación. En cada uno de los ejemplos, siempre existe un número infinito de opciones para cada vector característico. Escogeremos arbitrariamente una de ellas igualando una o más de las x_i a un número conveniente. Aquí hacemos una de las x_i igual a 1.

Ejemplo 5 Sea $A = \begin{pmatrix} 2 & -1 \\ -4 & 2 \end{pmatrix}$. Entonces $\det(A-\lambda I) = \begin{vmatrix} 2-\lambda & -1 \\ -4 & 2-\lambda \end{vmatrix} = \lambda^2 - 4\lambda = \lambda(\lambda-4)$. De esta manera los valores característicos son $\lambda_1 = 0$ y $\lambda_2 = 4$. El espacio característico correspondiente a cero es simplemente el núcleo o kernel de A. Calculamos $\begin{pmatrix} 2 & -1 \\ -4 & 2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ o $2x_1 = x_2$ y un vector característico es $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$. De esta manera $\ker A = E_0 = \text{gen}\left\{\begin{pmatrix} 1 \\ 2 \end{pmatrix}\right\}$. Correspondiente a $\lambda_2 = 4$ tenemos $\begin{pmatrix} -2 & -1 \\ -4 & -2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, por lo que $E_4 = \text{gen}\left\{\begin{pmatrix} 1 \\ -2 \end{pmatrix}\right\}$.

Ejemplo 6 Sea $A = \begin{pmatrix} 3 & -5 \\ 1 & -1 \end{pmatrix}$. Entonces $\det(A-\lambda I) = \begin{vmatrix} 3-\lambda & -5 \\ 1 & -1-\lambda \end{vmatrix} = \lambda^2 - 2\lambda + 2 = 0$

y

$$\lambda = \frac{-(-2) \pm \sqrt{4-4(1)(2)}}{2} = \frac{2 \pm \sqrt{-4}}{2} = \frac{2 \pm 2i}{2} = 1 \pm i$$

De esta manera $\lambda_1 = 1+i$ y $\lambda_2 = 1-i$. Calculamos

$$[A-(1+i)I]\mathbf{v} = \begin{pmatrix} 2-i & -5 \\ 1 & -2-i \end{pmatrix}^{*}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

---

* Note que las columnas de esta matriz son linealmente dependientes porque $\begin{pmatrix} -5 \\ -2-i \end{pmatrix} = (-2-i)\begin{pmatrix} 2-i \\ 1 \end{pmatrix}$.

y obtenemos $(2-i)x_1 - 5x_2 = 0$ y $x_1 + (-2-i)x_2 = 0$. De esta manera $x_1 = (2+i)x_2$, lo cual nos da el vector característico $\mathbf{v}_1 = \begin{pmatrix} 2+i \\ 1 \end{pmatrix}$ y $E_{1+i} = \text{gen}\left\{\begin{pmatrix} 2+i \\ 1 \end{pmatrix}\right\}$. Análogamente $[A-(1-i)I]\mathbf{v} = \begin{pmatrix} 2+i & -5 \\ 1 & -2+i \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ o bien $x_1 + (-2+i)x_2 = 0$, lo cual proporciona $x_1 = (2-i)x_2$, $\mathbf{v}_2 = \begin{pmatrix} 2-i \\ 1 \end{pmatrix}$, y $E_{1-i} = \text{gen}\left\{\begin{pmatrix} 2-i \\ 1 \end{pmatrix}\right\}$.

Observación 1. Este ejemplo muestra que una matriz real puede tener valores y vectores característicos complejos. Algunos textos definen los valores característicos de matrices reales como las raíces *reales* de la ecuación característica. Con esta definición la matriz del ejemplo anterior *no* tiene valores característicos. Esto puede simplificar los cálculos, pero también reduce enormemente la utilidad de la teoría de valores y vectores propios. Veremos una muestra significativa del uso de los valores propios complejos en la Sección 6.7.

Observación 2. Nótese que $\lambda_2 = 1 - i$ es el número complejo conjugado de $\lambda_1 = 1 + i$. También las componentes de $\mathbf{v}_2$ son las complejas conjugadas de $\mathbf{v}_1$. Esto no es coincidencia. En el Problema 33 se pide demostrar que

> los valores propios complejos de una matriz real siempre ocurren en pares conjugados
> y
> los correspondientes vectores propios son complejos conjugados entre sí.

Ejemplo 7 Sea $A = \begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix}$. Entonces $\det(A-\lambda I) = \begin{vmatrix} 4-\lambda & 0 \\ 0 & 4-\lambda \end{vmatrix} = (\lambda-4)^2 = 0$; por lo tanto $\lambda = 4$ es un valor característico de multiplicidad algebraica 2. Es obvio que $A\mathbf{v} = 4\mathbf{v}$ para todo vector $\mathbf{v} \in \mathbb{R}^2$, por lo que $E_4 = \mathbb{R}^2 = \text{gen}\left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right\}$.

Ejemplo 8 Sea $A = \begin{pmatrix} 4 & 1 \\ 0 & 4 \end{pmatrix}$. Entonces $\det(A-\lambda I) = \begin{vmatrix} 4-\lambda & 1 \\ 0 & 4-\lambda \end{vmatrix} = (\lambda-4)^2 = 0$; de esta manera $\lambda = 4$ es de nuevo un valor característico de multiplicidad algebraica 2. Pero ahora tenemos $(A-4I)\mathbf{v} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ 0 \end{pmatrix}$. En consecuencia, $x_2 = 0$, $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ es un vector característico, y $E_4 = \text{gen}\left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right\}$.

Ejemplo 9 Sea $A = \begin{pmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{pmatrix}$. Entonces $\det(A-\lambda I) = \begin{vmatrix} 3-\lambda & 2 & 4 \\ 2 & -\lambda & 2 \\ 4 & 2 & 3-\lambda \end{vmatrix} = -\lambda^3 + 6\lambda^2 + 15\lambda + 8 = -(\lambda+1)^2(\lambda-8) = 0$, por lo que los valores característicos son $\lambda_1 = 8$ y $\lambda_2 = -1$ (con multiplicidad algebraica 2). Para $\lambda_1 = 8$, obtenemos

$$(A-8I)\mathbf{v} = \begin{pmatrix} -5 & 2 & 4 \\ 2 & -8 & 2 \\ 4 & 2 & -5 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

o, reduciendo por renglones,

$$\left(\begin{array}{ccc|c} -5 & 2 & 4 & 0 \\ 2 & -8 & 2 & 0 \\ 4 & 2 & -5 & 0 \end{array}\right) \xrightarrow[A_{1,3}(-1)]{A_{1,2}(4)} \left(\begin{array}{ccc|c} -5 & 2 & 4 & 0 \\ -18 & 0 & 18 & 0 \\ 9 & 0 & -9 & 0 \end{array}\right) \xrightarrow{M_2(\frac{1}{18})}$$

$$\left(\begin{array}{ccc|c} -5 & 2 & 4 & 0 \\ -1 & 0 & 1 & 0 \\ 9 & 0 & -9 & 0 \end{array}\right) \xrightarrow{A_{2,3}(9)} \left(\begin{array}{ccc|c} 0 & 2 & -1 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Por lo tanto, $x_3 = 2x_2$ y $x_1 = x_3$, obtenemos el vector característico $\mathbf{v}_1 = \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}$ y $E_8 = \text{gen}\left\{\begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}\right\}$. Para $\lambda_2 = -1$, tenemos $(A+I)\mathbf{v} = \begin{pmatrix} 4 & 2 & 4 \\ 2 & 1 & 2 \\ 4 & 2 & 4 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$, lo cual proporciona únicamente la ecuación $2x_1 + x_2 + 2x_3 = 0$, o bien, $x_2 = -2x_1 - 2x_3$. Si $x_1 = 1$ y $x_3 = 0$, obtenemos $\mathbf{v}_2 = \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix}$. Si $x_1 = 0$ y $x_3 = 1$, entonces $\mathbf{v}_3 = \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix}$. De esta manera $E_{-1} = \text{gen}\left\{\begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix}\right\}$.

Existen otras opciones adecuadas para vectores característicos. Por ejemplo, $\mathbf{v} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$ está en E_{-1} ya que $\mathbf{v} = \mathbf{v}_2 - \mathbf{v}_3$.

Ejemplo 10 Sea $A = \begin{pmatrix} -5 & -5 & 9 \\ 8 & 9 & 18 \\ -2 & -3 & -7 \end{pmatrix}$. Entonces $\det(A-\lambda I) = \begin{vmatrix} -5-\lambda & -5 & -9 \\ 8 & 9-\lambda & 18 \\ -2 & -3 & -7-\lambda \end{vmatrix} = -\lambda^3 - 3\lambda^2 - 3\lambda - 1 = -(\lambda+1)^3 = 0$. De esta manera $\lambda = -1$ es un valor característico de multiplicidad algebraica 3. Para calcular E_{-1}, determinamos $(A+I)\mathbf{v} = \begin{pmatrix} -4 & -5 & -9 \\ 8 & 10 & 18 \\ -2 & -3 & -6 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ y reducimos por renglón para obtener, sucesivamente,

$$\left(\begin{array}{ccc|c} -4 & -5 & -9 & 0 \\ 8 & 10 & 18 & 0 \\ -2 & -3 & -6 & 0 \end{array}\right) \xrightarrow[A_{3,2}(4)]{A_{3,1}(-2)} \left(\begin{array}{ccc|c} 0 & 1 & 3 & 0 \\ 0 & -2 & -6 & 0 \\ -2 & -3 & -6 & 0 \end{array}\right) \xrightarrow[A_{1,3}(3)]{A_{1,2}(2)} \left(\begin{array}{ccc|c} 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 0 \\ -2 & 0 & 3 & 0 \end{array}\right)$$

Esto nos da $x_2 = -3x_3$ y $2x_1 = 3x_3$. Si hacemos $x_3 = 2$ obtenemos únicamente un

vector característico linealmente independiente: $\mathbf{v}_1=\begin{pmatrix}3\\-6\\2\end{pmatrix}$. De esta manera

$$E_{-1} = \text{gen}\left\{\begin{pmatrix}3\\-6\\2\end{pmatrix}\right\}.$$

Ejemplo 11 Sea $A=\begin{pmatrix}-1 & -3 & -9\\0 & 5 & 18\\0 & -2 & -7\end{pmatrix}$. Entonces $\det(A-\lambda I)=\begin{vmatrix}-1-\lambda & -3 & -9\\0 & 5-\lambda & 18\\0 & -2 & -7-\lambda\end{vmatrix}=$ $-(\lambda+1)^3=0$. De esta manera, como en el Ejemplo 10, $\lambda=-1$ es un valor característico de multiplicidad algebraica 3. Para encontrar E_{-1}, calculamos $(A+I)\mathbf{v}=\begin{pmatrix}0 & -3 & -9\\0 & 6 & 18\\0 & -2 & -6\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}$. Así $-2x_2-6x_3=0$ ó $x_2=-3x_3$, y x_1 es arbitrario. Haciendo $x_1=0$, $x_3=1$, obtenemos $\mathbf{v}_1=\begin{pmatrix}0\\-3\\1\end{pmatrix}$. Haciendo $x_1=1$, $x_3=1$ obtenemos $\mathbf{v}_2=\begin{pmatrix}1\\-3\\1\end{pmatrix}$. Por lo que $E_{-1}=\text{gen}\left\{\begin{pmatrix}0\\-3\\1\end{pmatrix},\begin{pmatrix}1\\-3\\1\end{pmatrix}\right\}$.

En cada uno de los ejemplos anteriores encontramos un valor característico con una multiplicidad algebraica de 2 o más. Pero como vimos en los Ejemplos 8, 10 y 11 el número de vectores característicos linealmente independientes no es necesariamente igual a la multiplicidad algebraica del valor característico (como era el caso en los Ejemplos 7 y 9). Esta observación nos conduce a la siguiente definición.

Definición 4 ***Multiplicidad geométrica.*** Sea λ un valor característico de A. Entonces la *multiplicidad geométrica* de λ es la dimensión del espacio característico correspondiente a λ (lo cual es la nulidad de la matriz $A-\lambda I$). Es decir:

$$\text{Multiplicidad geométrica de } \lambda = \dim E_\lambda = \nu(A-\lambda I)$$

En los Ejemplos 7 y 9 vimos que en los valores característicos de multiplicidad algebraica 2, la multiplicidad geométrica también era 2. En el Ejemplo 8 la multiplicidad geométrica de $\lambda=4$ era 1 mientras que la multiplicidad algebraica era 2. En el Ejemplo 10 la multiplicidad algebraica era 3 y la multiplicidad geométrica era 1. En el Ejemplo 11 la multiplicidad algebraica era 3 y la multiplicidad geométrica era 2. Estos ejemplos ilustran el hecho de que si la multiplicidad algebraica de λ es mayor que 1, entonces no podemos predecir la multiplicidad geométrica de λ sin información adicional.

Si A es una matriz de 2×2 y λ es un valor característico con multiplicidad algebraica 2, entonces la multiplicidad geométrica de λ es ≤ 2 puesto que pueden haber al menos dos vectores linealmente independientes en un espacio de dos dimensiones. Sea A una matriz de 3×3 con dos valores característicos λ_1 y λ_2 de multiplicidad algebraica 1 y 2, respectivamente. Entonces la multiplicidad geométrica de λ_2 es ≤ 2 porque de otra manera podríamos tener al menos cuatro vectores linealmente independientes en un espacio de tres dimensiones. Intuitivamente, parece que la multiplicidad geométrica de un valor característico es siempre menor o igual que la multiplicidad algebraica. La demostración del siguiente teorema no es difícil si se demuestran algunas propiedades adicionales sobre determinantes. Puesto que esto nos desviaría demasiado de nuestro objetivo, omitimos la demostración.*

Teorema 4 Sea λ un valor característico de A. Entonces:

$$\text{Multiplicidad geométrica de } \lambda \leq \text{ multiplicidad algebraica de } \lambda$$

Nota. La multiplicidad geométrica de un valor característico nunca es cero. Esto se establece de la Definición 1, la cual expresa que si λ es un valor característico, entonces existe un vector característico *diferente de cero* correspondiente a λ.

En el resto de este capítulo nos ocuparemos de determinar si una matriz dada de $n\times n$ tiene o no n vectores característicos linealmente independientes. De lo que hemos discutido hasta ahora, el siguiente teorema es evidente.

Teorema 5 Sea A una matriz de $n\times n$. Entonces A tiene n vectores característicos linealmente independientes si y sólo si la multiplicidad geométrica de todo valor característico es igual a la multiplicidad algebraica. En particular, A tiene n vectores característicos linealmente independientes si todos sus valores característicos son distintos (puesto que la multiplicidad algebraica de todo valor característico es 1).

En el Ejemplo 5, vimos una matriz para la cual el cero era un valor característico; en realidad, del Teorema 1 es claro que el cero es un valor característico de A si y sólo si $\det A=\det(A-0I)=0$. Esto nos permite ampliar, por última vez, el Teorema Resumen (vea Teorema 5.4.4).

---

* Vea la demostración del Teorema 11.2.6 en el libro de C.R. Wylie, *Advanced Engineering Mathematics* (Nueva York: McGraw-Hill, 1975).

Teorema 6 ***Teorema Resumen, Versión 8.*** Sea A una matriz de $n \times n$. Entonces los once enunciados siguientes son equivalentes. Esto es, si uno es cierto, todos son ciertos.

i. A es invertible.
ii. La única solución al sistema homogéneo $A\mathbf{x} = \mathbf{0}$ es la solución trivial $(\mathbf{x} = \mathbf{0})$.
iii. El sistema $A\mathbf{x} = \mathbf{b}$ tiene una solución única para todo n-vector $\mathbf{b}$.
iv. A es equivalente por renglón a la matriz identidad I_n de $n \times n$.
v. A puede escribirse como el producto de matrices elementales.
vi. Los renglones (y columnas) de A son linealmente independientes.
vii. $\det A \neq 0$.
viii. $\nu(A) = 0$.
ix. $\rho(A) = n$.
x. La transformación lineal T de $\mathbb{R}^n$ a $\mathbb{R}^n$ definida por $T\mathbf{x} = A\mathbf{x}$ es un isomorfismo.
xi. El cero *no* es un valor característico de A.

Problemas 6.1

En los Problemas del 1 *al* 20 *calcule los valores característicos y espacios característicos de la matriz dada. Si la multiplicidad algebraica de un valor característico es mayor que* 1, *calcule su multiplicidad geométrica.*

1. $\begin{pmatrix} -2 & -2 \\ -5 & 1 \end{pmatrix}$ **2.** $\begin{pmatrix} -12 & 7 \\ -7 & 2 \end{pmatrix}$ **3.** $\begin{pmatrix} 2 & -1 \\ 5 & -2 \end{pmatrix}$

4. $\begin{pmatrix} -3 & 0 \\ 0 & -3 \end{pmatrix}$ **5.** $\begin{pmatrix} -3 & 2 \\ 0 & -3 \end{pmatrix}$ **6.** $\begin{pmatrix} 3 & 2 \\ -5 & 1 \end{pmatrix}$

7. $\begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix}$ **8.** $\begin{pmatrix} 1 & 1 & -2 \\ -1 & 2 & 1 \\ 0 & 1 & -1 \end{pmatrix}$ **9.** $\begin{pmatrix} 5 & 4 & 2 \\ 4 & 5 & 2 \\ 2 & 2 & 2 \end{pmatrix}$

10. $\begin{pmatrix} 1 & 2 & 2 \\ 0 & 2 & 1 \\ -1 & 2 & 2 \end{pmatrix}$ **11.** $\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -3 & 3 \end{pmatrix}$ **12.** $\begin{pmatrix} -3 & -7 & -5 \\ 2 & 4 & 3 \\ 1 & 2 & 2 \end{pmatrix}$

13. $\begin{pmatrix} 1 & -1 & -1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix}$ **14.** $\begin{pmatrix} 7 & -2 & -4 \\ 3 & 0 & -2 \\ 6 & -2 & -3 \end{pmatrix}$ **15.** $\begin{pmatrix} 4 & 6 & 6 \\ 1 & 3 & 2 \\ -1 & -5 & -2 \end{pmatrix}$

16. $\begin{pmatrix} 4 & 1 & 0 & 1 \\ 2 & 3 & 0 & 1 \\ -2 & 1 & 2 & -3 \\ 2 & -1 & 0 & 5 \end{pmatrix}$ **17.** $\begin{pmatrix} a & 0 & 0 & 0 \\ 0 & a & 0 & 0 \\ 0 & 0 & a & 0 \\ 0 & 0 & 0 & a \end{pmatrix}$

18. $\begin{pmatrix} a & b & 0 & 0 \\ 0 & a & 0 & 0 \\ 0 & 0 & a & 0 \\ 0 & 0 & 0 & a \end{pmatrix}$; $b \neq 0$ **19.** $\begin{pmatrix} a & b & 0 & 0 \\ 0 & a & c & 0 \\ 0 & 0 & a & 0 \\ 0 & 0 & 0 & a \end{pmatrix}$; $bc \neq 0$

20. $\begin{pmatrix} a & b & 0 & 0 \\ 0 & a & c & 0 \\ 0 & 0 & a & d \\ 0 & 0 & 0 & a \end{pmatrix}$; $bcd \neq 0$

21. Muestre que para todo número real a y b, la matriz $A = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$ tiene los vectores característicos $\begin{pmatrix} 1 \\ i \end{pmatrix}$ y $\begin{pmatrix} 1 \\ -i \end{pmatrix}$.

En los Problemas del 22 *al* 28 *suponga que la matriz A tiene los valores característicos* $\lambda_1, \lambda_2, \ldots, \lambda_k$.

22. Muestre que los valores característicos de A^t son $\lambda_1, \lambda_2, \ldots, \lambda_k$.
23. Muestre que los valores característicos de αA son $\alpha\lambda_1, \alpha\lambda_2, \ldots, \alpha\lambda_k$.
24. Muestre que A^{-1} existe si y sólo si $\lambda_1\lambda_2 \cdots \lambda_k \neq 0$.
25. Si A^{-1} existe, muestre que los valores característicos de A^{-1} son $1/\lambda_1, 1/\lambda_2, \ldots, 1/\lambda_k$.
26. Muestre que la matriz $A - \alpha I$ tiene los valores característicos $\lambda_1 - \alpha, \lambda_2 - \alpha, \ldots, \lambda_k - \alpha$.
★ 27. Muestre que los valores característicos de A^2 son $\lambda_1^2, \lambda_2^2, \ldots, \lambda_k^2$.
★ 28. Muestre que los valores característicos de A^m son $\lambda_1^m, \lambda_2^m, \ldots, \lambda_k^m$ para $m = 1, 2, 3, \ldots$
29. Sea λ un valor característico de A con correspondiente vector característico $\mathbf{v}$. Sea $p(\lambda) = a_0 + a_1\lambda + a_2\lambda^2 + \cdots + a_n\lambda^n$. Defina la matriz $p(A)$ por $p(A) = a_0 I + a_1 A + a_2 A^2 + \cdots + a_n A^n$. Muestre que $p(A)\mathbf{v} = p(\lambda)\mathbf{v}$.
30. Usando el resultado del Problema 29, muestre que si $\lambda_1, \lambda_2, \ldots, \lambda_k$ son valores característicos de A, entonces $p(\lambda_1), p(\lambda_2), \ldots, p(\lambda_k)$, son valores característicos de $p(A)$.
31. Muestre que si A es una matriz triangular superior, entonces los valores característicos de A son las componentes de la diagonal de A.
32. Sea $A_1 = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$, $A_2 = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$, $A_3 = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$ y $A_4 = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}$. Muestre que, para cada matriz $\lambda = 2$ es un valor característico de multiplicidad algebraica 4. En cada caso, calcule la multiplicidad geométrica de $\lambda = 2$.
★ 33. Sea A una matriz real de $n \times n$. Muestre que si λ_1 es un valor característico complejo de A con vector característico $\mathbf{v}_1$, entonces $\bar{\lambda}_1$ es un valor característico de A con vector característico $\bar{\mathbf{v}}_1$.
★ 34. Una matriz de probabilidad es una matriz de $n \times n$ que tiene dos propiedades:
i. $a_{ij} \geq 0$ para toda i y j
ii. La suma de las componentes en cada columna es 1.
Demuestre que 1 es un valor característico para toda matriz de probabilidad.

6.2 Un modelo de crecimiento poblacional (opcional)

En esta sección mostraremos cómo la teoría de valores y vectores característicos puede utilizarse para analizar un modelo del crecimiento de una población de pájaros.* Comenzaremos por discutir un modelo simple de crecimiento poblacional. Supondremos que determinadas especies crecen en una proporción constante; esto es, la población de las especies después de un periodo (el cual podría ser una hora, una semana, un mes, un año, etc.) es un múltiplo constante de la población en el periodo anterior. Esto podría suceder, por ejemplo, si cada generación es distinta y cada organismo produce r descendientes y después muere. Si p_n denota la población después del n-ésimo periodo, tendríamos

$$p_n = rp_{n-1}$$

Por ejemplo, este modelo puede describir una población bacteriana donde, a un tiempo dado, un organismo se divide y origina dos organismos separados. Entonces $r = 2$. Sea p_0 la población inicial. Entonces $p_1 = rp_0$, $p_2 = rp_1 = r(rp_0) = r^2p_0$, $p_3 = rp_2 = r(r^2p_0) = r^3p_0$, y así sucesivamente, por lo que

$$p_n = r^n p_0 \tag{1}$$

De este modelo vemos que la población crece sin límites si $r > 1$ y decrece a cero si $r < 1$. Si $r = 1$ la población permanece a un nivel constante p_0.

Este modelo es, evidentemente, muy simple. Una objeción obvia es que el número de descendientes producidos depende, en muchos casos, de las edades de los adultos. Por ejemplo, en una población humana el promedio de mujeres adultas mayores de 50 años, produciría menos niños, que la mujer promedio de 21 años de edad. Para tratar esta dificultad, introduciremos un modelo que permita agrupar edades con diferentes tipos de fertilidad.

Ahora observaremos un modelo de crecimiento poblacional para especies de pájaros. En esta población de pájaros supondremos que el número de pájaros machos iguala al número de hembras. Denotemos por $p_{j,n-1}$ la población de hembras jóvenes (inmaduras) en el año $(n-1)$-ésimo y denotemos por $p_{a,n-1}$ el número de hembras adultas en el año $(n-1)$-ésimo. Algunos de los pájaros jóvenes morirán durante el año. Supondremos que una determinada proporción α de los pájaros jóvenes sobreviven para llegar a adultos en la primavera del n-ésimo año. Cada hembra sobreviviente produce huevos durante la primavera, los cuales se incuban para producir, a la siguiente primavera, en promedio, k pájaros hembras. Los adultos también mueren, y la proporción de adultos que sobreviven de una primavera a la siguiente es β.

Esta proporción constante de supervivencia de los pájaros no es una suposición simplificada. Esto sucede con la mayoría de las poblaciones naturales de pájaros que se han estudiado. Esto significa que la razón promedio de supervivencia de muchas especies de pájaros adultos es independiente de la edad. Quizás son pocos los pájaros que sobreviven lo suficiente para mostrar los

* El material de esta sección está basado en una publicación de D. Cooke: "A 2 × 2 Matrix Model of Population Growth", *Mathematical Gazette*, 61(416): 120-123.

efectos de la vejez. Más aún, en muchas especies el número de descendientes no parece estar influenciado por la edad de la madre.

En la notación introducida con anterioridad $p_{j,n}$ y $p_{a,n}$ representan, respectivamente, la población de hembras jóvenes y adultas en el año n-ésimo. Reuniendo toda la información dada, llegamos al siguiente sistema de 2×2:

$$\begin{aligned} p_{j,n} &= \qquad\qquad kp_{a,n-1} \\ p_{a,n} &= \alpha p_{j,n-1} + \beta p_{a,n-1} \end{aligned} \tag{2}$$

o bien

$$\mathbf{p}_n = A\mathbf{p}_{n-1} \tag{3}$$

donde $\mathbf{p}_n = \begin{pmatrix} p_{j,n} \\ p_{a,n} \end{pmatrix}$ y $A = \begin{pmatrix} 0 & k \\ \alpha & \beta \end{pmatrix}$. Resulta claro, de (3), que $\mathbf{p}_1 = A\mathbf{p}_0$, $\mathbf{p}_2 = A\mathbf{p}_1 = A(A\mathbf{p}_0) = A^2\mathbf{p}_0, \ldots$, y así sucesivamente. Por lo tanto

$$\boxed{\mathbf{p}_n = A^n\mathbf{p}_0} \tag{4}$$

donde $\mathbf{p}_0$ es el vector de las poblaciones iniciales de hembras jóvenes y adultas.

La Ecuación (4) es similar a la Ecuación (1), pero ahora podemos distinguir entre las razones de supervivencia de pájaros jóvenes y adultos.

Ejemplo 1 Sea $A = \begin{pmatrix} 0 & 2 \\ 0.3 & 0.5 \end{pmatrix}$. Esto significa que cada hembra adulta produce dos hembras jóvenes y, puesto que el número de machos se supone que es igual al número de hembras, al menos cuatro huevos (y probablemente muchos más, ya que la pérdida de pájaros jóvenes es grande). Del modelo es evidente que α y β se hallan en el intervalo $[0,1]$. Puesto que la mortalidad de los pájaros jóvenes es mayor que la de los adultos, debemos tener $\alpha < \beta$.

En la Tabla 6.1 suponemos que, inicialmente, existen 10 hembras y 10 machos adultos y que no existen pájaros jóvenes. Los cálculos fueron hechos

Tabla 6.1

Año n	No. de jóvenes $p_{j,n}$	No. de adultos $p_{a,n}$	Población total de hembras T_n en el n-ésimo año	$p_{j,n}/p_{a,n}$*	T_n/T_{n-1}*
0	0	10	10	0	—
1	20	5	25	4.00	2.50
2	10	8	18	1.18	0.74
3	17	7	24	2.34	1.31
4	14	8	22	1.66	0.96
5	17	8	25	2.00	1.13
10	22	12	34	1.87	1.06
11	24	12	36	1.88	1.07
12	25	13	38	1.88	1.06
20	42	22	64	1.88	1.06

* Las cifras en estas columnas fueron obtenidas antes que las de las columnas previas fueran redondeadas. De esta manera, por ejemplo, en el año 2, $p_{j,2}/p_{a,2} = 10/8.5 = 1.176470588 \approx 1.18$.

en una computadora, pero el trabajo no es tan complicado y puede hacerse con una calculadora de bolsillo. Por ejemplo, $\mathbf{p}_1 = \begin{pmatrix} 0 & 2 \\ 0.3 & 0.5 \end{pmatrix}\begin{pmatrix} 0 \\ 10 \end{pmatrix} = \begin{pmatrix} 20 \\ 5 \end{pmatrix}$, por lo que $p_{j,1} = 20$, $p_{a,1} = 5$, la población total de hembras después de un año es de 25, y la proporción de hembras jóvenes a adultas es de 4 a 1. En el segundo año, $\mathbf{p}_2 = \begin{pmatrix} 0 & 2 \\ 0.3 & 0.5 \end{pmatrix}\begin{pmatrix} 20 \\ 5 \end{pmatrix} = \begin{pmatrix} 10 \\ 8.5 \end{pmatrix}$, el cual redondeamos a $\begin{pmatrix} 10 \\ 8 \end{pmatrix}$ ya que no podemos tener $8\frac{1}{2}$ pájaros adultos. En la Tabla 6.1 se presentan las razones $p_{j,n}/p_{a,n}$ y las razones T_n/T_{n-1} del número total de hembras en años sucesivos.

En la Tabla 6.1 parece como si la razón $p_{j,n}/p_{a,n}$ se aproximara a la constante 1.88 mientras que la población total parece estar incrementándose a una razón constante de 6 por ciento anual. Veamos si podemos determinar por qué sucede esto.

Primero, regresemos al caso general (Ecuación 4). Suponga que A tiene los valores característicos reales λ_1 y λ_2 diferentes con vectores característicos correspondientes $\mathbf{v}_1$ y $\mathbf{v}_2$. Puesto que $\mathbf{v}_1$ y $\mathbf{v}_2$ son linealmente independientes, podemos escribir

$$\mathbf{p}_0 = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 \tag{5}$$

para algunos números reales a_1 y a_2. Entonces (4) se convierte en

$$\mathbf{p}_n = A^n(a_1\mathbf{v}_1 + a_2\mathbf{v}_2) \tag{6}$$

Pero $A\mathbf{v}_1 = \lambda_1\mathbf{v}_1$ y $A^2\mathbf{v}_1 = A(A\mathbf{v}_1) = A(\lambda_1\mathbf{v}_1) = \lambda_1 A\mathbf{v}_1 = \lambda_1(\lambda_1\mathbf{v}_1) = \lambda_1^2\mathbf{v}_1$. De esta manera podemos ver que $A^n\mathbf{v}_1 = \lambda_1^n\mathbf{v}_1$, $A^n\mathbf{v}_2 = \lambda_2^n\mathbf{v}_2$, y, de (6),

$$\mathbf{p}_n = a_1\lambda_1^n\mathbf{v}_1 + a_2\lambda_2^n\mathbf{v}_2 \tag{7}$$

La ecuación característica de A es $\begin{vmatrix} -\lambda & k \\ \alpha & \beta-\lambda \end{vmatrix} = \lambda^2 - \beta\lambda - k\alpha = 0$ o $\lambda = (\beta \pm \sqrt{\beta^2+4\alpha k})/2$. Por hipótesis, $k>0$, $0<\alpha<1$ y $0<\beta<1$. Por lo tanto $4\alpha k>0$ y $\beta^2+4\alpha k>0$, lo cual significa que los valores característicos son, en efecto, reales y distintos y que un valor característico λ_1 es positivo, el otro λ_2 es negativo y $|\lambda_1|>|\lambda_2|$. Podemos escribir (7) como

$$\mathbf{p}_n = \lambda_1^n\left[a_1\mathbf{v}_1 + \left(\frac{\lambda_2}{\lambda_1}\right)^n a_2\mathbf{v}_2\right] \tag{8}$$

Puesto que $|\lambda_2/\lambda_1|<1$, es evidente que $(\lambda_2/\lambda_1)^n$ se hace muy pequeño conforme n crece. De esta manera, para n grande

$$\mathbf{p}_n \approx a_1\lambda_1^n\mathbf{v}_1 \tag{9}$$

Esto significa que a la larga la distribución de edad se estabiliza y es proporcional a $\mathbf{v}_1$. Cada grupo de edad cambiará cada año por un factor de λ_1. De esta manera, a la larga, la Ecuación (4) actúa exactamente como la Ecuación (1). A corto plazo, es decir, antes de que se alcance la "estabilidad", los números oscilan. La magnitud de esta oscilación depende de la magnitud de λ_2/λ_1 (lo cual es negativo, explicando de esta manera la oscilación).

Ejemplo 1 (continuación) Para $A = \begin{pmatrix} 0 & 2 \\ 0.3 & 0.5 \end{pmatrix}$, tenemos $\lambda^2 - 0.5\lambda - 0.6 = 0$ o $\lambda = (0.5 \pm \sqrt{0.25+2.4})/2 = (0.5 \pm \sqrt{2.65})/2$, por lo que $\lambda_1 \approx 1.06$ y $\lambda_2 \approx -0.56$. Esto explica el incremento de 6 por ciento en la población que notamos en la última columna de la Tabla 6.1. Para el valor característico $\lambda_1 = 1.06$, calculamos $(A - 1.06I)\mathbf{v}_1 = \begin{pmatrix} -1.06 & 2 \\ 0.3 & -0.56 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ o $1.06x_1 = 2x_2$, por lo que $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0.53 \end{pmatrix}$ es un vector característico. Análogamente $(A + 0.56)\mathbf{v}_2 = \begin{pmatrix} 0.56 & 2 \\ 0.3 & 1.06 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, por lo que $0.56x_1 + 2x_2 = 0$ y $\mathbf{v}_2 = \begin{pmatrix} 1 \\ -0.28 \end{pmatrix}$ es un segundo vector característico. Note que en $\mathbf{v}_1$, tenemos $1/0.53 \approx 1.88$. Esto explica la razón $p_{j,n}/p_{a,n}$ en la quinta columna de la tabla.

Observación. En los cálculos anteriores la precisión se perdió ya que redondeamos a dos decimales. Mayor exactitud se obtiene usando una calculadora o una computadora. Por ejemplo, usando una calculadora de bolsillo fácilmente podemos calcular $\lambda_1 = 1.06394103$, $\lambda_2 = -0.5639410298$, $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0.531970515 \end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix} 1 \\ -0.2819705149 \end{pmatrix}$, así pues, la razón de $p_{j,n}$ a $p_{a,n}$ es $1/0.531970515 \approx 1.879801537$.

Es importante hacer notar cuánta información se obtiene de un simple cálculo de valores característicos. Es de gran interés conocer si una población en última instancia crecerá o decrecerá. Aumentará si $\lambda_1>1$, y la condición para ello es $(\beta+\sqrt{\beta^2+4\alpha k})/2>1$ o $\sqrt{\beta^2+4\alpha k}>2-\beta$ o $\beta^2+4\alpha k>(2-\beta)^2=4-4\beta+\beta^2$. Esto conduce a $4\alpha k>4-4\beta$ o bien

$$k > \frac{1-\beta}{\alpha} \tag{10}$$

En el Ejemplo 1 teníamos $\beta=0.5$, $\alpha=0.3$; de esta manera (10) se satisface si $k>0.5/0.3\approx1.67$.

Indicamos dos limitaciones de este modelo antes de concluir con esta sección:

i. Las proporciones de nacimientos y muertes a menudo cambian de año a año y son dependientes, en particular, del clima. Este modelo presupone un medio ambiente constante.

ii. Los ecólogos han descubierto que para muchas especies la proporción de nacimientos y muertes varía con el tamaño de la población. En particular, una población no puede crecer cuando alcanza un determinado tamaño de-

bido a la limitación de alimentos y espacio. Es obvio que una población no puede crecer indefinidamente a una proporción constante. De otra manera esa población podría inundar la tierra.

Problemas 6.2

En los Problemas del 1 *al* 3 *encuentre el número de pájaros jóvenes y de hembras adultas después de* 1, 2, 5, 10, 19 *y* 20 *años. Después encuentre las razones, a la larga, de* $p_{j,n}$ *a* $p_{a,n}$ *y* T_n *a* T_{n-1}. [Sugerencia: *Use las Ecuaciones* (7) *y* (9), *una calculadora y redondee a tres decimales.*]

1. $\mathbf{p}_0=\begin{pmatrix}0\\12\end{pmatrix}$; $k=3$, $\alpha=0.4$, $\beta=0.6$
2. $\mathbf{p}_0=\begin{pmatrix}0\\15\end{pmatrix}$; $k=1$, $\alpha=0.3$, $\beta=0.4$
3. $\mathbf{p}_0=\begin{pmatrix}0\\20\end{pmatrix}$; $k=4$, $\alpha=0.7$, $\beta=0.8$
4. Muestre que si $\alpha=\beta$ y $\alpha>\frac{1}{2}$, entonces la población de pájaros a la larga crecerá, si por lo menos nace una hembra en promedio por cada hembra adulta.
5. Muestre que a la larga $p_{j,n}/p_{a,n}$ se aproxima al valor límite k/λ_1.
6. Suponga que dividimos los pájaros adultos en dos grupos de edades: aquéllos de 1 a 5 años de edad y aquéllos de más de 5 años de edad. Suponga que la proporción de supervivencia en el primer grupo es β, mientras que en el segundo grupo es γ (y $\beta>\gamma$). Suponga que los pájaros del primer grupo están divididos en grupos de igual número por edades. (Esto es, si hay 100 pájaros en el grupo, entonces 20 tienen un año de edad, 20 tienen 2 años de edad, y así sucesivamente.) Formule una matriz modelo de 3×3 para este caso.

6.3 Matrices equivalentes y diagonalización

En esta sección vamos a describir una relación útil e interesante que es válida entre dos matrices.

Definición 1 **Matrices equivalentes.** Dos matrices A y B de $n\times n$ se llaman *equivalentes* (o *semejantes*) si existe una matriz invertible C de $n\times n$ tal que

$$B=C^{-1}AC \qquad (1)$$

La función definida por (1) asocia a la matriz A la matriz B y se llama *transformación de equivalencia*.

Nota. $C^{-1}(A_1+A_2)C=C^{-1}A_1C+C^{-1}A_2C$ y $C^{-1}(\alpha A)C=\alpha C^{-1}AC$, por lo

que la función definida por (1) es, de hecho, una transformación lineal. Esto explica el uso de la palabra "transformación" en la Definición 1.

El propósito de esta sección es mostrar que (i) las matrices equivalentes tiene varias propiedades importantes en común y (ii) que la mayor parte de las matrices son equivalentes a matrices diagonales.

Ejemplo 1 Sean $A=\begin{pmatrix}2&1\\0&-1\end{pmatrix}$, $B=\begin{pmatrix}4&-2\\5&-3\end{pmatrix}$, y $C=\begin{pmatrix}2&-1\\-1&1\end{pmatrix}$. Entonces $CB=\begin{pmatrix}2&-1\\-1&1\end{pmatrix}\begin{pmatrix}4&-2\\5&-3\end{pmatrix}=\begin{pmatrix}3&-1\\1&-1\end{pmatrix}$ y $AC=\begin{pmatrix}2&1\\0&-1\end{pmatrix}\begin{pmatrix}2&-1\\-1&1\end{pmatrix}=\begin{pmatrix}3&-1\\1&-1\end{pmatrix}$. De esta manera $CB=AC$. Puesto que $\det C=1\neq 0$, C es invertible; ya que $CB=AC$, tenemos $C^{-1}CB=C^{-1}AC$ o bien $B=C^{-1}AC$. Esto muestra que A y B son equivalentes.

Ejemplo 2 Sean $D=\begin{pmatrix}1&0&0\\0&-1&0\\0&0&2\end{pmatrix}$, $A=\begin{pmatrix}-6&-3&-25\\2&1&8\\2&2&7\end{pmatrix}$ y $C=\begin{pmatrix}2&4&3\\0&1&-1\\3&5&7\end{pmatrix}$. C es invertible porque $\det C=3\neq 0$. Entonces calculamos:

$$CA=\begin{pmatrix}2&4&3\\0&1&-1\\3&5&7\end{pmatrix}\begin{pmatrix}-6&-3&-25\\2&1&8\\2&2&7\end{pmatrix}=\begin{pmatrix}2&4&3\\0&-1&1\\6&10&14\end{pmatrix}$$

$$DC=\begin{pmatrix}1&0&0\\0&-1&0\\0&0&2\end{pmatrix}\begin{pmatrix}2&4&3\\0&1&-1\\3&5&7\end{pmatrix}=\begin{pmatrix}2&4&3\\0&-1&1\\6&10&14\end{pmatrix}$$

De esta manera $CA=DC$ y $A=C^{-1}DC$, por lo que A y D son equivalentes.

Nota. En los Ejemplos 1 y 2 no fue necesario calcular C^{-1}. Únicamente era necesario conocer que C es no singular.

Teorema 1 Si A y B son matrices equivalentes de $n\times n$, entonces A y B tienen la misma ecuación característica y, por consiguiente tienen los mismos valores característicos.

Demostración Puesto que A y B son equivalentes, $B=C^{-1}AC$ y

$$\begin{aligned}\det(B-\lambda I)&=\det(C^{-1}AC-\lambda I)=\det[C^{-1}AC-C^{-1}(\lambda I)C]\\&=\det[C^{-1}(A-\lambda I)C]=\det(C^{-1})\det(A-\lambda I)\det(C)\\&=\det(C^{-1})\det(C)\det(A-\lambda I)=\det(C^{-1}C)\det(A-\lambda I)\\&=\det I\det(A-\lambda I)=\det(A-\lambda I)\end{aligned}$$

Esto significa que A y B tienen la misma ecuación característica y, puesto que los valores característicos son raíces de la ecuación característica, también tienen los mismos valores característicos. ■

Ejemplo 3 En el Ejemplo 2 es obvio que los valores característicos de $D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$ son 1, −1 y 2. De esta manera estos son los valores característicos de $A = \begin{pmatrix} -6 & -3 & -25 \\ 2 & 1 & 8 \\ 2 & 2 & 7 \end{pmatrix}$. Compruebe esto verificando que $\det(A - I) = \det(A + I) = \det(A - 2I) = 0$.

En muchas aplicaciones es de considerable utilidad "diagonalizar" una matriz A, esto es, encontrar una matriz diagonal equivalente a A.

Definición 2 **Matriz diagonalizable.** Una matriz A de $n \times n$ es *diagonalizable* si existe una matriz diagonal D tal que A es equivalente a D.

Observación. Si D es una matriz diagonal, entonces sus valores característicos son las componentes de su diagonal. Si A es equivalente a D tenemos los mismos valores característicos (Teorema 1). Si juntamos estos dos resultados, observamos que si A es diagonalizable, entonces A es equivalente a una matriz diagonal y las componentes de la diagonal son los valores característicos de A.

El siguiente teorema nos dice cuándo una matriz es diagonalizable.

Teorema 2 Una matriz A de $n \times n$ es diagonalizable si y sólo si tiene n vectores característicos linealmente independientes. En este caso, la matriz diagonal D equivalente a A está dada por

$$D = \begin{pmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{pmatrix} \qquad (2)$$

donde $\lambda_1, \lambda_2, \ldots, \lambda_n$ son los valores característicos de A. Si C es una matriz cuyas columnas son vectores característicos linealmente independientes de A, entonces

$$D = C^{-1}AC \qquad (3)$$

Demostración Supondremos primero que A tiene n vectores característicos linealmente independientes $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ correspondientes a los valores característicos $\lambda_1, \lambda_2, \ldots, \lambda_n$ (no necesariamente diferentes).

Sean

$$\mathbf{v}_1 = \begin{pmatrix} c_{11} \\ c_{21} \\ \vdots \\ c_{n1} \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} c_{12} \\ c_{22} \\ \vdots \\ c_{n2} \end{pmatrix}, \ldots, \quad \mathbf{v}_n = \begin{pmatrix} c_{1n} \\ c_{2n} \\ \vdots \\ c_{nn} \end{pmatrix}$$

y sea

$$C = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{pmatrix}$$

Entonces C es invertible puesto que sus columnas son linealmente independientes. Ahora

$$AC = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{pmatrix}$$

y vemos que la i-ésima columna de AC es $A\begin{pmatrix} c_{1i} \\ c_{2i} \\ \vdots \\ c_{ni} \end{pmatrix} = A\mathbf{v}_i = \lambda_i \mathbf{v}_i$. De esta manera AC es la matriz cuya columna i-ésima es $\lambda_i \mathbf{v}_i$ y

$$AC = \begin{pmatrix} \lambda_1 c_{11} & \lambda_2 c_{12} & \cdots & \lambda_n c_{1n} \\ \lambda_1 c_{21} & \lambda_2 c_{22} & \cdots & \lambda_n c_{2n} \\ \vdots & \vdots & & \vdots \\ \lambda_1 c_{n1} & \lambda_2 c_{n2} & \cdots & \lambda_n c_{nn} \end{pmatrix}$$

Pero

$$CD = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} = \begin{pmatrix} \lambda_1 c_{11} & \lambda_2 c_{12} & \cdots & \lambda_n c_{1n} \\ \lambda_1 c_{21} & \lambda_2 c_{22} & \cdots & \lambda_n c_{2n} \\ \vdots & \vdots & & \vdots \\ \lambda_1 c_{n1} & \lambda_2 c_{n2} & \cdots & \lambda_n c_{nn} \end{pmatrix}$$

De esta manera

$$AC = CD \tag{4}$$

y, puesto que C es invertible, podemos multiplicar ambos lados de (4) por la izquierda por C^{-1} para obtener

$$D = C^{-1}AC \tag{5}$$

Esto demuestra que si A tiene n vectores característicos linealmente independientes, entonces A es diagonalizable. Inversamente, suponga que A es diagonalizable. Esto es, suponga que (5) es válida para alguna matriz invertible C. Sean $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ las columnas de C. Entonces $AC = CD$ e invirtiendo los argumentos anteriores, inmediatamente vemos que $A\mathbf{v}_i = \lambda_i \mathbf{v}_i$ para $i = 1, 2, \ldots, n$. De esta manera $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ son vectores característicos de A y son linealmente independientes ya que C es invertible. ■

Notación Para indicar que D es una matriz diagonal con componentes diagonales $\lambda_1, \lambda_2, \ldots, \lambda_n$, escribimos $D = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$.

El Teorema 2 tiene un corolario útil que se obtiene inmediatamente del Teorema 6.1.3.

Corolario Si la matriz A de $n \times n$ tiene n valores característicos distintos, entonces A es diagonalizable.

Observación. Si los coeficientes reales de un polinomio de grado n, son tomados al azar, entonces, con probabilidad 1, el polinomio tendrá n raíces diferentes. Intuitivamente, no es difícil ver por qué sucede esto. Si $n = 2$, por ejemplo, entonces la ecuación $\lambda^2 + a\lambda + b = 0$ tiene raíces iguales si y sólo si $a^2 = 4b$ —un caso muy improbable si a y b se escogen al azar. Podemos escribir, por supuesto, polinomios que tengan raíces de multiplicidad algebraica mayor que 1, pero estos polinomios son excepcionales. De modo que, sin tratar de ser matemáticamente preciso, es válido decir que la *mayor parte* de los polinomios tienen raíces diferentes. Por lo tanto la *mayor parte* de las matrices tienen valores característicos diferentes y, por lo visto al principio de la sección, la *mayor parte* de las matrices son diagonalizables.

Ejemplo 4 Sea $A = \begin{pmatrix} 4 & 2 \\ 3 & 3 \end{pmatrix}$. En el Ejemplo 6.1.3 encontramos los dos vectores característicos linealmente independientes $\mathbf{v}_1 = \begin{pmatrix} 2 \\ -3 \end{pmatrix}$ y $\mathbf{v}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Entonces si hacemos $C = \begin{pmatrix} 2 & 1 \\ -3 & 1 \end{pmatrix}$, encontramos que

$$C^{-1}AC = \frac{1}{5}\begin{pmatrix} 1 & -1 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} 4 & 2 \\ 3 & 3 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ -3 & 1 \end{pmatrix}$$

$$= \frac{1}{5}\begin{pmatrix} 1 & -1 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} 2 & 6 \\ -3 & 6 \end{pmatrix} = \frac{1}{5}\begin{pmatrix} 5 & 0 \\ 0 & 30 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 6 \end{pmatrix}$$

que es la matriz cuyos componentes diagonales son los valores característicos de A.

Ejemplo 5 Sea $A = \begin{pmatrix} 1 & -1 & 4 \\ 3 & 2 & -1 \\ 2 & 1 & -1 \end{pmatrix}$. En el Ejemplo 6.1.4 calculamos los tres vectores característicos linealmente independientes $\mathbf{v}_1 = \begin{pmatrix} -1 \\ 4 \\ 1 \end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}$ y $\mathbf{v}_3 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$. Entonces $C = \begin{pmatrix} -1 & 1 & 1 \\ 4 & -1 & 2 \\ 1 & -1 & 1 \end{pmatrix}$ y

$$C^{-1}AC = -\frac{1}{6}\begin{pmatrix} 1 & -2 & 3 \\ -2 & -2 & 6 \\ -3 & 0 & -3 \end{pmatrix}\begin{pmatrix} 1 & -1 & 4 \\ 3 & 2 & -1 \\ 2 & 1 & -1 \end{pmatrix}\begin{pmatrix} -1 & 1 & 1 \\ 4 & -1 & 2 \\ 1 & -1 & 1 \end{pmatrix}$$

$$= -\frac{1}{6}\begin{pmatrix} 1 & -2 & 3 \\ -2 & -2 & 6 \\ -3 & 0 & -3 \end{pmatrix}\begin{pmatrix} -1 & -2 & 3 \\ 4 & 2 & 6 \\ 1 & 2 & 3 \end{pmatrix}$$

$$= -\frac{1}{6}\begin{pmatrix} -6 & 0 & 0 \\ 0 & 12 & 0 \\ 0 & 0 & -18 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

con valores característicos 1, -2 y 3.

Observación. Puesto que existe un número infinito de formas de escoger un vector característico, existe un número infinito de formas de escoger la matriz diagonalizante C. La única recomendación es escoger los vectores característicos y la matriz C tal que simplifiquen los cálculos aritméticos. Esto significa que se deberán incluir tantos ceros y unos como sea posible.

Ejemplo 6 Sea $A = \begin{pmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{pmatrix}$. Entonces del Ejemplo 6.1.9, tenemos los tres vectores característicos linealmente independientes $\mathbf{v}_1 = \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix}$ y $\mathbf{v}_3 = \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix}$. Haciendo $C = \begin{pmatrix} 2 & 1 & 0 \\ 1 & -2 & -2 \\ 2 & 0 & 1 \end{pmatrix}$, obtenemos

$$C^{-1}AC = -\frac{1}{9}\begin{pmatrix} -2 & -1 & -2 \\ -5 & 2 & 4 \\ 4 & 2 & -5 \end{pmatrix}\begin{pmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{pmatrix}\begin{pmatrix} 2 & 1 & 0 \\ 1 & -2 & -2 \\ 2 & 0 & 1 \end{pmatrix}$$

$$= -\frac{1}{9}\begin{pmatrix} -2 & -1 & -2 \\ -5 & 2 & 4 \\ 4 & 2 & -5 \end{pmatrix}\begin{pmatrix} 16 & -1 & 0 \\ 8 & 2 & 2 \\ 16 & 0 & -1 \end{pmatrix}$$

$$= -\frac{1}{9}\begin{pmatrix} -72 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 9 \end{pmatrix} = \begin{pmatrix} 8 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

Este ejemplo ilustra que A es diagonalizable aun cuando sus valores característicos no son diferentes.

Ejemplo 7 Sea $A = \begin{pmatrix} 4 & 1 \\ 0 & 4 \end{pmatrix}$. En el Ejemplo 6.1.8 vimos que A *no* tenía dos vectores característicos linealmente independientes. Suponga que A fuera diagonalizable (en contradicción con el Teorema 2). Entonces $D = \begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix}$ y habría una matriz invertible C tal que $C^{-1}AC = D$. Multiplicando esta ecuación por la izquierda por C y por la derecha por C^{-1}, resulta entonces que $A = CDC^{-1} = C\begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix}C^{-1} = C(4I)C^{-1} = 4CIC^{-1} = 4CC^{-1} = 4I = \begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix} = D$. Pero $A \neq D$, por lo que C no existe.

Hemos visto que muchas matrices son equivalentes a matrices diagonales. Sin embargo, todavía tenemos dos preguntas:

i. ¿Es posible determinar si una matriz dada es diagonalizable sin calcular los valores y vectores característicos?
ii. ¿Qué haremos si A no es diagonalizable?

Podemos encontrar una respuesta parcial a la primera pregunta en la siguiente sección y una respuesta completa a la segunda en la Sección 6.6. En la Sección 6.7 podemos ver una aplicación importante del procedimiento de diagonalización.

Al principio de este capítulo definimos los vectores propios y los valores propios de una transformación lineal $T: V \to V$, donde $\dim V = n$. Dijimos entonces que debido a que T puede ser representada por una matriz $n \times n$, limitaríamos nuestra discusión al caso de las matrices $n \times n$.

Sin embargo, una transformación lineal puede ser representada por muchas matrices $n \times n$, una para cada elección de base. ¿Tienen todas estas matrices representantes los mismos valores propios y vectores propios? La respuesta es sí, como se demuestra en el siguiente teorema.

Teorema 3 Sea V un espacio finito-dimensional con bases $B_1 = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \nu_n\}$ y $B_2 = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n\}$. Sea $T: V \to V$ una transformación lineal. Si A_T es la representación matricial de T con respecto a la base B_1, y C_T es la representación matricial de T respecto de la base B_2, entonces A_T y C_T son semejantes.

Demostración T es una transformación lineal de V en sí mismo. Del Teorema 5.3.3

$$(T\mathbf{x})_{B_1} = A_T(\mathbf{x})_{B_1} \tag{6}$$

y

$$(T\mathbf{x})_{B_2} = C_T(\mathbf{x})_{B_2} \tag{7}$$

Sea M la matriz de transición de B_1 a B_2. Entonces, por el Teorema 4.9.1,

$$(\mathbf{x})_{B_2} = M(\mathbf{x})_{B_1} \tag{8}$$

para toda $\mathbf{x}$ en V. También

$$(T\mathbf{x})_{B_2} = M(T\mathbf{x})_{B_1} \tag{9}$$

Introduciendo (8) y (9) en (7),

$$M(T\mathbf{x})_{B_1} = C_T M(\mathbf{x})_{B_1} \tag{10}$$

La matriz M es invertible por el resultado del Teorema 4.9.2. Si multiplicamos ambos lados de (10) por M^{-1} (que es la matriz de transición de B_2 a B_1), entonces

$$(T\mathbf{x})_{B_1} = M^{-1}C_T M(\mathbf{x})_{B_1} \tag{11}$$

Comparando (6) con (11) tenemos que

$$A_T(\mathbf{x})_{B_1} = M^{-1}C_TM(\mathbf{x})_{B_1} \tag{12}$$

Como (12) es válida para todo $\mathbf{x} \in V$, se concluye que

$$A_T = M^{-1}C_TM$$

Esto es, A_T y C_T son semejantes o equivalentes. ■

Problemas 6.3

En los Problemas del 1 *al* 15 *determine si la matriz dada A es diagonalizable. Si lo es, encuentre una matriz C tal que* $C^{-1}AC = D$.

1. $\begin{pmatrix} -2 & -2 \\ -5 & 1 \end{pmatrix}$ 2. $\begin{pmatrix} 3 & -1 \\ -2 & 4 \end{pmatrix}$ 3. $\begin{pmatrix} 2 & -1 \\ 5 & -2 \end{pmatrix}$

4. $\begin{pmatrix} 3 & -5 \\ 1 & -1 \end{pmatrix}$ 5. $\begin{pmatrix} 3 & 2 \\ -5 & 1 \end{pmatrix}$ 6. $\begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix}$

7. $\begin{pmatrix} 1 & 1 & -2 \\ -1 & 2 & 1 \\ 0 & 1 & -1 \end{pmatrix}$ 8. $\begin{pmatrix} 2 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ 9. $\begin{pmatrix} 3 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 2 \end{pmatrix}$

10. $\begin{pmatrix} 3 & -1 & -1 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \end{pmatrix}$ 11. $\begin{pmatrix} 7 & -2 & -4 \\ 3 & 0 & -2 \\ 6 & -2 & -3 \end{pmatrix}$ 12. $\begin{pmatrix} 4 & 6 & 6 \\ 1 & 3 & 2 \\ -1 & -5 & -2 \end{pmatrix}$

13. $\begin{pmatrix} -3 & -7 & -5 \\ 2 & 4 & 3 \\ 1 & 2 & 2 \end{pmatrix}$ 14. $\begin{pmatrix} -2 & -2 & 0 & 0 \\ -5 & 1 & 0 & 0 \\ 0 & 0 & 2 & -1 \\ 0 & 0 & 5 & -2 \end{pmatrix}$ 15. $\begin{pmatrix} 4 & 1 & 0 & 1 \\ 2 & 3 & 0 & 1 \\ -2 & 1 & 2 & -3 \\ 2 & -1 & 0 & 5 \end{pmatrix}$

16. Muestre que si A es equivalente a B y B es equivalente a C, entonces A es equivalente a C.
17. Si A es semejante a B, demuestre que A^n es semejante a B^n para todo número entero positivo n.
★ 18. Si A es equivalente a B, muestre que $\rho(A)=\rho(B)$ y $\nu(A)=\nu(B)$. [*Sugerencia:* Primero demuestre que si C es invertible, entonces $\nu(CA)=\nu(A)$ mostrando que $\mathbf{x}\in N_A$ si y sólo si $\mathbf{x}\in N_{CA}$. En seguida demuestre que $\rho(AC)=\rho(A)$ mostrando que $R_A = R_{AC}$. Concluya que $\rho(AC) = \rho(CA) = \rho(A)$. Finalmente, haga uso del hecho de que C^{-1} es invertible para mostrar que $\rho(C^{-1}AC) = \rho(A)$.]
19. Sea $D=\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Calcule D^{20}.
20. Si A es equivalente a B muestre que $\det A = \det B$.

21. Suponga que $C^{-1}AC = D$. Muestre que para todo entero n, $A^n = CD^nC^{-1}$. Esto da un modo fácil de calcular potencias de una matriz diagonalizable.
22. Sea $A=\begin{pmatrix} 3 & -4 \\ 2 & -3 \end{pmatrix}$. Calcule A^{20}. [*Sugerencia:* Halle una C tal que $A = CDC^{-1}$.]
★ 23. Sea A una matriz de $n \times n$ cuya ecuación característica es $(\lambda - c)^n = 0$. Muestre que A es diagonalizable si y sólo si $A = cI$.
24. Use el resultado del Problema 21 y el Ejemplo 6 para calcular A^{10}, donde $A = \begin{pmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{pmatrix}$.
★ 25. Sean A y B matrices reales de $n \times n$ con valores característicos distintos. Demuestre que $AB = BA$ si y sólo si A y B tienen los mismos vectores característicos.
26. Si A es diagonalizable muestre que $\det A = \lambda_1, \lambda_2, \ldots, \lambda_n$, donde $\lambda_1, \lambda_2, \ldots, \lambda_n$ son los valores característicos de A.

6.4 Matrices simétricas y diagonalización ortogonal

En esta sección veremos que las matrices simétricas* tienen una cantidad de propiedades importantes. En particular, mostraremos que cualquier matriz simétrica tiene n vectores característicos reales linealmente independientes y por consiguiente, por el Teorema 6.3.2, es diagonalizable. Empezaremos por demostrar que los valores característicos de una matriz simétrica real son reales.

Teorema 1 Sea A una matriz simétrica real de $n \times n$. Entonces los valores característicos de A son reales.

Demostración ‡ Sea λ un valor característico de A con vector característico $\mathbf{v}$; esto es $A\mathbf{v}=\lambda\mathbf{v}$. Ahora $\mathbf{v}$ es un vector en $\mathbb{C}^n$, y un producto interno en $\mathbb{C}^n$ (vea Definición 4.11.1 y Ejemplo 4.11.2) satisface

$$(\alpha\mathbf{x}, \mathbf{y}) = \alpha(\mathbf{x}, \mathbf{y}) \quad \text{y} \quad (\mathbf{x}, \alpha\mathbf{y}) = \bar{\alpha}(\mathbf{x}, \mathbf{y}) \tag{1}$$

Entonces

$$(A\mathbf{v}, \mathbf{v}) = (\lambda\mathbf{v}, \mathbf{v}) = \lambda(\mathbf{v}, \mathbf{v}) \tag{2}$$

Más aún, por el Teorema 5.5.1 y el hecho de que $A^t = A$,

$$(A\mathbf{v}, \mathbf{v}) = (\mathbf{v}, A^t\mathbf{v}) = (\mathbf{v}, A\mathbf{v}) = (\mathbf{v}, \lambda\mathbf{v}) = \bar{\lambda}(\mathbf{v}, \mathbf{v}) \tag{3}$$

De esta manera, igualando (2) y (3) tenemos

$$\lambda(\mathbf{v}, \mathbf{v}) = \bar{\lambda}(\mathbf{v}, \mathbf{v}) \tag{4}$$

Pero $(\mathbf{v}, \mathbf{v}) = |\mathbf{v}|^2 \neq 0$ puesto que $\mathbf{v}$ es un vector característico. De esta manera podemos dividir ambos lados de (4) entre $(\mathbf{v}, \mathbf{v})$ para obtener

* Recuerde que A es simétrica si y sólo si $A^t = A$.
‡ Esta demostración utiliza material de las Secciones 4.11 y 5.5 y podría ser omitida si esas secciones no se han visto.

$$\lambda = \bar{\lambda} \tag{5}$$

Si $\lambda = a + ib$, entonces $\bar{\lambda} = a - ib$ y, de (5), tenemos

$$a + ib = a - ib \tag{6}$$

lo cual puede ser válido únicamente si $b = 0$. Esto muestra que $\lambda = a$; por lo tanto λ es real y el teorema está demostrado. ■

Vimos en el Teorema 6.1.3 que vectores característicos correspondientes a valores característicos diferentes son linealmente independientes. Para matrices simétricas el resultado es más categórico: *los vectores característicos de una matriz simétrica correspondientes a valores característicos diferentes son ortogonales.*

Teorema 2 Sea A una matriz simétrica real de $n \times n$. Si λ_1 y λ_2 son valores característicos distintos con correspondientes vectores característicos reales $\mathbf{v}_1$ y $\mathbf{v}_2$, entonces $\mathbf{v}_1$ y $\mathbf{v}_2$ son ortogonales.

Demostración Calculamos

$$A\mathbf{v}_1 \cdot \mathbf{v}_2 = \lambda_1 \mathbf{v}_1 \cdot \mathbf{v}_2 = \lambda_1(\mathbf{v}_1 \cdot \mathbf{v}_2) \tag{7}$$

y

$$A\mathbf{v}_1 \cdot \mathbf{v}_2 = \mathbf{v}_1 \cdot A^t\mathbf{v}_2 = \mathbf{v}_1 \cdot A\mathbf{v}_2 = \mathbf{v}_1 \cdot (\lambda_2\mathbf{v}_2) = \lambda_2(\mathbf{v}_1 \cdot \mathbf{v}_2) \tag{8}$$

Combinando (7) y (8) tenemos $\lambda_1(\mathbf{v}_1 \cdot \mathbf{v}_2) = \lambda_2(\mathbf{v}_1 \cdot \mathbf{v}_2)$ y puesto que $\lambda_1 \neq \lambda_2$, concluimos que $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$. Esto es lo que queríamos demostrar. ■

Ahora formularemos el resultado principal de esta sección. Su demostración es difícil (y optativa) y aparece al final de esta sección.

Teorema 3 Sea A una matriz simétrica real de $n \times n$. Resulta entonces que A tiene n vectores característicos reales ortonormales.

Observación. Se concluye de este teorema que la multiplicidad geométrica de cada valor característico de A es igual a su multiplicidad algebraica.

El Teorema 3 nos dice que si A es simétrica entonces $\mathbb{R}^n$ tiene una base $B = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ que consiste de vectores característicos ortonormales de A. Sea Q la matriz cuyas columnas son $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$. Entonces por el Teorema 4.10.3, Q es una matriz ortogonal. Esto nos conduce hacia la siguiente definición.

Definición 1 **Matriz ortogonalmente diagonalizable.** Una matriz A de $n \times n$ se dice que es *diagonalizable ortogonalmente* si existe una matriz ortogonal Q tal que

$$Q^tAQ = D \tag{9}$$

donde $D = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ y $\lambda_1, \lambda_2, \ldots, \lambda_n$ son los valores característicos de A.

Nota. Recuerde que Q es ortogonal si $Q^t = Q^{-1}$; por lo tanto (9) podría ser escrita como $Q^{-1}AQ = D$.

Teorema 4 Sea A una matriz real de $n \times n$. Entonces A es diagonalizable ortogonalmente si y sólo si A es simétrica.

Demostración Sea A simétrica. Entonces, por los Teoremas 2 y 3, A es diagonalizable ortogonalmente con Q (la matriz cuyas columnas son los vectores característicos ortonormales dados en el Teorema 3). Inversamente, suponga que A es diagonalizable ortogonalmente. Entonces existe una matriz ortogonal Q tal que $Q^tAQ = D$. Multiplicando esta ecuación por la izquierda por Q, por la derecha por Q^t y utilizando el hecho de que $Q^tQ = QQ^t = I$, obtenemos

$$A = QDQ^t \tag{10}$$

Entonces $A^t = (QDQ^t)^t = (Q^t)^tD^tQ^t = QDQ^t = A$. De esta manera A es simétrica y el teorema está demostrado. En la última serie de ecuaciones usamos los hechos de que $(AB)^t = B^tA^t$ (parte *ii* del Teorema 1.9.1) y $(A^t)^t = A$ (parte *i* del Teorema 1.9.1) y $D^t = D$ para cualquier matriz diagonal D. ■

Antes de dar ejemplos, enunciaremos los tres pasos que sirven para encontrar la matriz ortogonal Q que diagonaliza a la matriz simétrica A:

Procedimiento para encontrar una matriz diagonalizante Q:

i. Encuentre una base para cada espacio característico de A.
ii. Encuentre una base ortonormal para cada espacio característico de A, usando el procedimiento Gram-Schmidt.
iii. Establezca a Q como la matriz cuyas columnas son los vectores característicos ortonormales obtenidos en el paso (*ii*).

Ejemplo 1 Sea $A = \begin{pmatrix} 1 & -2 \\ -2 & 3 \end{pmatrix}$. Entonces la ecuación característica de A es $\det(A - \lambda I) = \begin{vmatrix} 1-\lambda & -2 \\ -2 & 3-\lambda \end{vmatrix} = \lambda^2 - 4\lambda - 1 = 0$, ecuación que tiene las raíces $\lambda = (4 \pm \sqrt{20})/2 = (4 \pm 2\sqrt{5})/2 = 2 \pm \sqrt{5}$. En el caso de $\lambda_1 = 2 - \sqrt{5}$, resulta así que $(A - \lambda I)\mathbf{v} = \begin{pmatrix} -1+\sqrt{5} & -2 \\ -2 & 1+\sqrt{5} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$. Un vector característico es $\mathbf{v}_1 = \begin{pmatrix} 2 \\ -1+\sqrt{5} \end{pmatrix}$, $|\mathbf{v}_1| = \sqrt{2^2 + (-1+\sqrt{5})^2} = \sqrt{10 - 2\sqrt{5}}$. De esta manera $\mathbf{u}_1 = 1/(\sqrt{10-2\sqrt{5}})\begin{pmatrix} 2 \\ -1+\sqrt{5} \end{pmatrix}$.

Para $\lambda_2 = 2 + \sqrt{5}$, calculamos $(A-\lambda I)\mathbf{v} = \begin{pmatrix} -1-\sqrt{5} & -2 \\ -2 & 1-\sqrt{5} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ y $\mathbf{v}_2 = \begin{pmatrix} 1-\sqrt{5} \\ 2 \end{pmatrix}$. Note que $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$ (lo cual debe ser cierto de conformidad con el Teorema 2). Se tiene así que $|\mathbf{v}_2| = \sqrt{10-2\sqrt{5}}$, por lo que $\mathbf{u}_2 = (1/\sqrt{10-2\sqrt{5}})\begin{pmatrix} 1-\sqrt{5} \\ 2 \end{pmatrix}$. Finalmente,

$$Q = \frac{1}{\sqrt{10-2\sqrt{5}}}\begin{pmatrix} 2 & 1-\sqrt{5} \\ -1+\sqrt{5} & 2 \end{pmatrix}, \quad Q^t = \frac{1}{\sqrt{10-2\sqrt{5}}}\begin{pmatrix} 2 & -1+\sqrt{5} \\ 1-\sqrt{5} & 2 \end{pmatrix}$$

y

$$Q^tAQ = \frac{1}{10-2\sqrt{5}}\begin{pmatrix} 2 & -1+\sqrt{5} \\ 1-\sqrt{5} & 2 \end{pmatrix}\begin{pmatrix} 1 & -2 \\ -2 & 3 \end{pmatrix}\begin{pmatrix} 2 & 1-\sqrt{5} \\ -1+\sqrt{5} & 2 \end{pmatrix}$$

$$= \frac{1}{10-2\sqrt{5}}\begin{pmatrix} 2 & -1+\sqrt{5} \\ 1-\sqrt{5} & 2 \end{pmatrix}\begin{pmatrix} 4-2\sqrt{5} & -3-\sqrt{5} \\ -7+3\sqrt{5} & 4+2\sqrt{5} \end{pmatrix}$$

$$= \frac{1}{10-2\sqrt{5}}\begin{pmatrix} 30-14\sqrt{5} & 0 \\ 0 & 10+6\sqrt{5} \end{pmatrix} = \begin{pmatrix} 2-\sqrt{5} & 0 \\ 0 & 2+\sqrt{5} \end{pmatrix}$$

Ejemplo 2 Sea $A = \begin{pmatrix} 5 & 4 & 2 \\ 4 & 5 & 2 \\ 2 & 2 & 2 \end{pmatrix}$. Entonces A resulta ser simétrica y $\det(A-\lambda I) = \begin{vmatrix} 5-\lambda & 4 & 2 \\ 4 & 5-\lambda & 2 \\ 2 & 2 & 2-\lambda \end{vmatrix} = -(\lambda-1)^2(\lambda-10)$. Para $\lambda = 1$ calculamos los vectores característicos linealmente independientes $\mathbf{v}_1 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$ y $\mathbf{v}_2 = \begin{pmatrix} -1 \\ 0 \\ 2 \end{pmatrix}$. Para $\lambda = 10$, encontramos que $\mathbf{v}_3 = \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}$. Para encontrar Q aplicamos el proceso de Gram-Schmidt a $\{\mathbf{v}_1, \mathbf{v}_2\}$, que es una base de E_1. Puesto que $|\mathbf{v}_1| = \sqrt{2}$, hacemos $\mathbf{u}_1 = \begin{pmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix}$. Después

$$\mathbf{v}_2' = \mathbf{v}_2 - (\mathbf{v}_2 \cdot \mathbf{u}_1)\mathbf{u}_1 = \begin{pmatrix} -1 \\ 0 \\ 2 \end{pmatrix} - \frac{1}{\sqrt{2}}\begin{pmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ 2 \end{pmatrix} - \begin{pmatrix} -1/2 \\ 1/2 \\ 0 \end{pmatrix} = \begin{pmatrix} -1/2 \\ -1/2 \\ 2 \end{pmatrix}$$

Entonces $|\mathbf{v}_2'| = \sqrt{18/4} = 3\sqrt{2}/2$ y $\mathbf{u}_2 = \dfrac{2}{3\sqrt{2}}\begin{pmatrix} -1/2 \\ -1/2 \\ 2 \end{pmatrix} = \begin{pmatrix} -1/3\sqrt{2} \\ -1/3\sqrt{2} \\ 4/3\sqrt{2} \end{pmatrix}$. Comprobamos esto notando que $\mathbf{u}_1 \cdot \mathbf{u}_2 = 0$. Finalmente, tenemos $\mathbf{u}_3 = \mathbf{v}_3/|\mathbf{v}_3| = \frac{1}{3}\mathbf{v}_3 = \begin{pmatrix} 2/3 \\ 2/3 \\ 1/3 \end{pmatrix}$.

Esto lo podemos comprobar también notando que $\mathbf{u}_1 \cdot \mathbf{u}_3 = 0$ y $\mathbf{u}_2 \cdot \mathbf{u}_3 = 0$. Por lo tanto

$$Q = \begin{pmatrix} -1/\sqrt{2} & -1/3\sqrt{2} & 2/3 \\ 1/\sqrt{2} & -1/3\sqrt{2} & 2/3 \\ 0 & 4/3\sqrt{2} & 1/3 \end{pmatrix}$$

y

$$Q^tAQ = \begin{pmatrix} -1/\sqrt{2} & 1/\sqrt{2} & 0 \\ -1/3\sqrt{2} & -1/3\sqrt{2} & 4/3\sqrt{2} \\ 2/3 & 2/3 & 1/3 \end{pmatrix}\begin{pmatrix} 5 & 4 & 2 \\ 4 & 5 & 2 \\ 2 & 2 & 2 \end{pmatrix}\begin{pmatrix} -1/\sqrt{2} & -1/3\sqrt{2} & 2/3 \\ 1/\sqrt{2} & -1/3\sqrt{2} & 2/3 \\ 0 & 4/3\sqrt{2} & 1/3 \end{pmatrix}$$

$$= \begin{pmatrix} -1/\sqrt{2} & 1/\sqrt{2} & 0 \\ -1/3\sqrt{2} & -1/3\sqrt{2} & 4/3\sqrt{2} \\ 2/3 & 2/3 & 1/3 \end{pmatrix}\begin{pmatrix} -1/\sqrt{2} & -1/3\sqrt{2} & 20/3 \\ 1/\sqrt{2} & -1/3\sqrt{2} & 20/3 \\ 0 & 4/3\sqrt{2} & 10/3 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 10 \end{pmatrix}$$

En esta sección hemos demostrado resultados para matrices simétricas reales. Si $A = (a_{ij})$ es una matriz compleja, entonces la *transpuesta conjugada* de A, denotada por A^* está definida por: el elemento ij-ésimo de $A^* = \overline{a_{ji}}$. La matriz A es llamada *hermitiana* si $A^* = A$. Resulta que los Teoremas 1, 2 y 3 también son ciertos para matrices hermitianas. Más aún, si definimos una matriz *unitaria* como una matriz compleja U con $U^* = U^{-1}$, entonces, usando la demostración del Teorema 4, podemos mostrar que una matriz hermitiana es diagonalizable unitariamente. Dejamos todos estos resultados como ejercicios (vea Problemas 15 a 17).

Concluimos esta sección con una demostración del Teorema 3.

Demostración del Teorema 3* Vamos a demostrar que a cada valor característico λ de multiplicidad algebraica k, le corresponden k vectores característicos ortonormales. Este paso, combinado con el Teorema 2 será suficiente. Sea $\mathbf{u}_1$ un vector característico de A correspondiente a λ_1. Podemos suponer que $|\mathbf{u}_1| = 1$. También podemos suponer que $\mathbf{u}_1$ es real ya que λ_1 es real y $\mathbf{u}_1 \in N_{A-\lambda_1 I}$, el núcleo de la matriz real $A - \lambda_1 I$. Este núcleo es un subespacio de $\mathbb{R}^n$, por el Ejemplo 4.6.10. En seguida notamos que $\{\mathbf{u}_1\}$ puede expandirse a una base $\{\mathbf{u}_1, \mathbf{v}_2, \mathbf{v}_3, \ldots, \mathbf{v}_n\}$ para $\mathbb{R}^n$ y, por el proceso de Gram-Schmidt, podemos transformar esta base en la base ortonormal $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$. Sea Q la matriz ortogonal cuyas columnas

* Si el tiempo lo permite.

son $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$. Por conveniencia de notación se establece que $Q = (\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n)$. Ahora Q es invertible y $Q^t = Q^{-1}$, por lo que A es equivalente a Q^tAQ y, por el Teorema 6.3.1, Q^tAQ y A tienen el mismo polinomio característico: $|Q^tAQ - \lambda I| = |A - \lambda I|$. Por consiguiente,

$$Q^t = \begin{pmatrix} \mathbf{u}_1^t \\ \mathbf{u}_2^t \\ \vdots \\ \mathbf{u}_n^t \end{pmatrix}$$

por lo que

$$Q^tAQ = \begin{pmatrix} \mathbf{u}_1^t \\ \mathbf{u}_2^t \\ \vdots \\ \mathbf{u}_n^t \end{pmatrix} A(\mathbf{u}_1\mathbf{u}_2 \cdots \mathbf{u}_n) = \begin{pmatrix} \mathbf{u}_1^t \\ \mathbf{u}_2^t \\ \vdots \\ \mathbf{u}_n^t \end{pmatrix} (A\mathbf{u}_1 A\mathbf{u}_2 \cdots A\mathbf{u}_n)$$

$$= \begin{pmatrix} \mathbf{u}_1^t \\ \mathbf{u}_2^t \\ \vdots \\ \mathbf{u}_n^t \end{pmatrix} (\lambda_1\mathbf{u}_1, A\mathbf{u}_2 \cdots A\mathbf{u}_n) \begin{pmatrix} \lambda_1 & \mathbf{u}_1^tA\mathbf{u}_2 & \cdots & \mathbf{u}_1^tA\mathbf{u}_n \\ 0 & \mathbf{u}_2^tA\mathbf{u}_2 & \cdots & \mathbf{u}_2^tA\mathbf{u}_n \\ \vdots & \vdots & & \vdots \\ 0 & \mathbf{u}_n^tA\mathbf{u}_2 & \cdots & \mathbf{u}_n^tA\mathbf{u}_n \end{pmatrix}$$

Los ceros aparecen porque $\mathbf{u}_1^t\mathbf{u}_j = \mathbf{u}_1 \cdot \mathbf{u}_j = 0$ si $j \neq 1$. Ahora bien $[Q^tAQ]^t = Q^tA^t(Q^t)^t = Q^tAQ$. En consecuencia, Q^tAQ es simétrica, lo cual significa que en el primer renglón de Q^tAQ deben existir ceros, para igualar los ceros en la primera columna.

De esta manera,

$$Q^tAQ = \begin{pmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & q_{22} & q_{23} & \cdots & q_{2n} \\ 0 & q_{32} & q_{33} & \cdots & q_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & q_{n2} & q_{n3} & \cdots & q_{nn} \end{pmatrix}$$

y

$$|Q^tAQ - \lambda I| = \begin{vmatrix} \lambda_1 - \lambda & 0 & 0 & \cdots & 0 \\ 0 & q_{22}-\lambda & q_{23} & \cdots & q_{2n} \\ 0 & q_{32} & q_{33}-\lambda & \cdots & q_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & q_{n2} & q_{n3} & \cdots & q_{nn}-\lambda \end{vmatrix}$$

$$= (\lambda_1 - \lambda) \begin{vmatrix} q_{22}-\lambda & q_{23} & \cdots & q_{2n} \\ q_{32} & q_{33}-\lambda & \cdots & q_{3n} \\ \vdots & \vdots & & \vdots \\ q_{n2} & q_{n3} & \cdots & q_{nn}-\lambda \end{vmatrix} = (\lambda - \lambda_1)\,|M_{11}(\lambda)|$$

donde $M_{11}(\lambda)$ es el menor 1, 1 de $Q^tAQ - \lambda I$. Si $k = 1$, no hay nada que demostrar. Si $k > 1$, entonces $|A - \lambda I|$ tiene el factor $(\lambda - \lambda_1)^2$ y, por lo tanto, $|Q^tAQ - \lambda I|$ también contiene el factor $(\lambda - \lambda_1)^2$. Consecuentemente, $|M_{11}(\lambda)|$ contiene el factor $\lambda - \lambda_1$, lo cual significa que $|M_{11}(\lambda_1)| = 0$. Esto quiere decir que las últimas $n - 1$ columnas de $Q^tAQ - \lambda_1 I$ son linealmente dependientes. Puesto que la primera columna de $Q^tAQ - \lambda_1 I$ es el vector cero, esto significa que $Q^tAQ - \lambda_1 I$ contiene a lo sumo $n - 2$ columnas linealmente independientes. En otras palabras, $\rho(Q^tAQ - \lambda_1 I) \leq n - 2$. Pero $Q^tAQ - \lambda_1 I$ y $A - \lambda_1 I$ son equivalentes; de aquí que por el Problema 6.3.18, $\rho(A - \lambda_1 I) \leq n - 2$. Por ello, $\nu(A - \lambda_1 I) \geq 2$, lo cual quiere decir que E_λ = núcleo de $(A - \lambda_1 I)$ contiene al menos dos vectores característicos linealmente independientes. Si $k = 2$ el resultado es claro. Si $k > 2$, entonces tomamos dos vectores ortonormales $\mathbf{u}_1, \mathbf{u}_2$ en E_λ y los expandimos a una nueva base ortonormal $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ para $\mathbb{R}^n$ y definimos $P = (\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n)$. Entonces, exactamente como antes mostramos que

$$P^tAP - \lambda I = \begin{pmatrix} \lambda_1-\lambda & 0 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_1-\lambda & 0 & 0 & \cdots & 0 \\ 0 & 0 & \beta_{33}-\lambda & \beta_{34} & \cdots & \beta_{3n} \\ 0 & 0 & \beta_{43} & \beta_{44}-\lambda & \cdots & \beta_{4n} \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \beta_{n3} & \beta_{n4} & \cdots & \beta_{nn}-\lambda \end{pmatrix}$$

Debido a que $k > 2$, mostramos, como antes, que el determinante de la matriz entre corchetes es cero cuando $\lambda = \lambda_1$ lo cual muestra que $\rho(P^tAP - \lambda_1 I) \leq n - 3$, por lo que $\nu(P^tAP - \lambda_1 I) = \nu(A - \lambda_1 I) \geq 3$. Entonces $\dim E_{\lambda_1} \geq 3$ y así sucesivamente. Podemos, claramente, continuar este proceso para mostrar que $\dim E_{\lambda_1} = k$. Para terminar, en cada E_{λ_i} podemos encontrar una base ortonormal. Esto completa la demostración. ■

Problemas 6.4

En los Problemas del 1 al 8 encuentre una matriz ortogonal que diagonalice la matriz simétrica dada.

1. $\begin{pmatrix} 3 & 4 \\ 4 & -3 \end{pmatrix}$ **2.** $\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ **3.** $\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$

4. $\begin{pmatrix} 1 & -1 & -1 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \end{pmatrix}$ 5. $\begin{pmatrix} -1 & 2 & 2 \\ 2 & -1 & 2 \\ 2 & 2 & 1 \end{pmatrix}$ 6. $\begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix}$

7. $\begin{pmatrix} 3 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 4 \end{pmatrix}$ 8. $\begin{pmatrix} 1 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$

9. Sea Q una matriz ortogonal simétrica. Muestre que si λ es un valor característico de Q, entonces $\lambda = \pm 1$.

10. A es *equivalente ortogonalmente* a B si existe una matriz ortogonal Q tal que $B = Q^tAQ$. Suponga que A es equivalente ortogonalmente a B y que B es equivalente ortogonalmente a C. Muestre que A es equivalente ortogonalmente a C.

11. Muestre que $Q = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ es ortogonal entonces $b = \pm c$. [*Sugerencia:* Escriba las ecuaciones que resulten de la ecuación $Q^tQ = I$.]

12. Suponga que A es una matriz simétrica real con cada uno de sus valores característicos igual a cero. Muestre que A es la matriz cero.

13. Muestre que si una matriz real A de 2×2 tiene vectores característicos que son ortogonales, entonces A es simétrica.

14. Sea A una matriz antisimétrica ($A^t = -A$). Demuestre que cada valor característico de A es de la forma $i\alpha$, donde α es un número real. Esto es, demuestre que cada valor característico de A es un número *imaginario puro*.

★ 15. Muestre que los valores característicos de una matriz hermitiana compleja de $n \times n$ son reales. [*Sugerencia:* Use el hecho de que en $\mathbb{C}^n$, $(A\mathbf{x}, \mathbf{y}) = (\mathbf{x}, A^*\mathbf{y})$.]

★ 16. Si A es una matriz hermitiana de $n \times n$, muestre que los vectores característicos correspondientes a valores característicos diferentes son ortogonales.

★★ 17. Repitiendo la demostración del Teorema 3, excepto que $\bar{\mathbf{v}}_i^t$ sustituye $\mathbf{v}_i^t$ donde sea adecuado, muestre que cualquier matriz hermitiana de $n \times n$ tiene n vectores característicos ortonormales.

18. Encuentre una matriz unitaria U tal que U^*AU sea diagonal, donde $A = \begin{pmatrix} 1 & 1-i \\ 1+i & 0 \end{pmatrix}$.

19. Haga lo mismo para $A = \begin{pmatrix} 2 & 3-3i \\ 3+3i & 5 \end{pmatrix}$.

20. Demuestre que el determinante de una matriz hermitiana es real.

6.5 Formas cuadráticas y secciones cónicas

En esta sección usaremos el material de la Sección 6.4 para encontrar información acerca de las gráficas de ecuaciones cuadráticas. Las ecuaciones cuadráticas y las formas cuadráticas, que se definen más adelante, surgen de diversas maneras. Por ejemplo, podemos usar formas cuadráticas para obtener información acerca de las secciones cónicas en $\mathbb{R}^2$ (circunferencias, parábolas, elipses, hipérbolas) y extender esta teoría para describir ciertas superficies llamadas *superficies cuádricas* en $\mathbb{R}^3$. Estos temas serán discutidos más adelante en esta sección. Aunque no lo discutiremos en este texto, las formas cuadráticas

surgen en un gran número de aplicaciones que van desde una descripción de funciones de costos en economía hasta el análisis del control de un cohete que viaja en el espacio.

Definición 1 **Ecuación cuadrática y forma cuadrática.**

i. Una *ecuación cuadrática en dos variables con términos no lineales* es una ecuación de la forma

$$ax^2 + bxy + cy^2 = d \qquad (1)$$

donde $|a| + |b| + |c| \neq 0$.

ii. Una *forma cuadrática en dos variables* es una expresión de la forma

$$F(x, y) = ax^2 + bxy + cy^2 \qquad (2)$$

donde $|a| + |b| + |c| \neq 0$.

Obviamente las ecuaciones cuadráticas y formas cuadráticas están íntimamente relacionadas. Empezaremos nuestro análisis de formas cuadráticas con un ejemplo sencillo.

Considere la forma cuadrática $F(x, y) = x^2 - 4xy + 3y^2$. Sea $\mathbf{v} = \begin{pmatrix} x \\ y \end{pmatrix}$ y $A = \begin{pmatrix} 1 & -2 \\ -2 & 3 \end{pmatrix}$. Entonces

$$A\mathbf{v}\cdot\mathbf{v} = \begin{pmatrix} 1 & -2 \\ -2 & 3 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}\cdot\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x-2y \\ -2x+3y \end{pmatrix}\cdot\begin{pmatrix} x \\ y \end{pmatrix}$$
$$= (x^2 - 2xy) + (-2xy + 3y^2) = x^2 - 4xy + 3y^2 = F(x, y)$$

De esta manera hemos "representado" la forma cuadrática $F(x,y)$ por la matriz simétrica A en el sentido de que

$$F(x, y) = A\mathbf{v}\cdot\mathbf{v} \qquad (3)$$

Recíprocamente, si A es una matriz simétrica, entonces la Ecuación (3) define una forma cuadrática $F(x, y) = A\mathbf{v}\cdot\mathbf{v}$.

La función $F(x,y)$ puede ser representada por muchas matrices pero sólo por una matriz simétrica. Para ver esto, sea $A = \begin{pmatrix} 1 & a \\ b & 3 \end{pmatrix}$, donde $a + b = -4$. Entonces $A\mathbf{v}\cdot\mathbf{v} = F(x, y)$. Si $A = \begin{pmatrix} 1 & 3 \\ -7 & 3 \end{pmatrix}$, por ejemplo, entonces $A\mathbf{v} =$

$\begin{pmatrix} x+3y \\ -7x+3y \end{pmatrix}$ y $A\mathbf{v} \cdot \mathbf{v} = x^2 - 4xy + 3y^2$. Sin embargo, si insistimos en que A sea simétrica, entonces debemos tener $a + b = -4$ y $a = b$. Este par de ecuaciones tienen la solución única $a = b = -2$.

Si $F(x, y) = ax^2 + bxy + cy^2$ es una forma cuadrática, sea

$$A = \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix} \tag{4}$$

Entonces

$$A\mathbf{v}\cdot\mathbf{v} = \left[\begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}\right]\cdot\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax+(b/2)y \\ (b/2)x+cy \end{pmatrix}\cdot\begin{pmatrix} x \\ y \end{pmatrix}$$
$$= ax^2 + bxy + cy^2 = F(x, y)$$

Ahora regresemos a la Ecuación cuadrática (1). Usando (3), podemos escribir (1) como

$$A\mathbf{v}\cdot\mathbf{v} = d \tag{5}$$

donde A es simétrica. Por el Teorema 6.4.4, existe una matriz ortogonal Q tal que $Q^tAQ = D$, donde $D = \text{diag}(\lambda_1, \lambda_2)$ y λ_1 y λ_2 son los valores característicos de A. Entonces $A = QDQ^t$ (recuerde que $Q^t = Q^{-1}$) y (5) puede ser escrito

$$(QDQ^t\mathbf{v})\cdot\mathbf{v} = d \tag{6}$$

Pero, por el Teorema 5.5.1, $A\mathbf{v} \cdot \mathbf{y} = \mathbf{v} \cdot A^t\mathbf{y}$. De esta manera

$$Q(DQ^t\mathbf{v})\cdot\mathbf{v} = DQ^t\mathbf{v}\cdot Q^t\mathbf{v} \tag{7}$$

por lo que (6) se puede escribir

$$[DQ^t\mathbf{v}]\cdot Q^t\mathbf{v} = d \tag{8}$$

Sea $\mathbf{v}' = Q^t\mathbf{v}$. Entonces $\mathbf{v}'$ es un vector de dimensión 2 y (8) se convierte en

$$D\mathbf{v}'\cdot\mathbf{v}' = d \tag{9}$$

Veamos (9) más de cerca. Podemos escribir $\mathbf{v}' = \begin{pmatrix} x' \\ y' \end{pmatrix}$. Ya que una matriz diagonal es simétrica, (9) define una forma cuadrática $F'(x', y')$ en las variables x' y y'. Si $D = \begin{pmatrix} a' & 0 \\ 0 & c' \end{pmatrix}$, entonces $D\mathbf{v}' = \begin{pmatrix} a' & 0 \\ 0 & c' \end{pmatrix}\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a'x' \\ c'y' \end{pmatrix}$ y

$$F'(x', y') = D\mathbf{v}'\cdot\mathbf{v}' = \begin{pmatrix} a'x' \\ c'y' \end{pmatrix}\cdot\begin{pmatrix} x' \\ y' \end{pmatrix} = a'x'^2 + c'y'^2$$

Esto es: $F'(x', y')$ *es una forma cuadrática sin el término* $x'y'$. Por lo tanto la Ecuación (9) es una ecuación cuadrática en las nuevas variables $x'y'$ donde el término $x'y'$ ha desaparecido.

Ejemplo 1 Considere la ecuación cuadrática $x^2 - 4xy + 3y^2 = 6$. Entonces, como hemos visto, la ecuación puede escribirse en la forma $A\mathbf{x}\cdot\mathbf{x} = 6$, donde $A = \begin{pmatrix} 1 & -2 \\ -2 & 3 \end{pmatrix}$. Anteriormente, en el Ejemplo 6.4.1, vimos que A puede ser diagonalizada a $D = \begin{pmatrix} 2-\sqrt{5} & 0 \\ 0 & 2+\sqrt{5} \end{pmatrix}$ usando la matriz ortogonal

$$Q = \frac{1}{\sqrt{10-2\sqrt{5}}}\begin{pmatrix} 2 & 1-\sqrt{5} \\ -1+\sqrt{5} & 2 \end{pmatrix}$$

Entonces
$$\mathbf{x}' = \begin{pmatrix} x' \\ y' \end{pmatrix} = Q^t\mathbf{x} = \frac{1}{\sqrt{10-2\sqrt{5}}}\begin{pmatrix} 2 & -1+\sqrt{5} \\ 1-\sqrt{5} & 2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$$
$$= \frac{1}{\sqrt{10-2\sqrt{5}}}\begin{pmatrix} 2x+(-1+\sqrt{5})y \\ (1-\sqrt{5})x+2y \end{pmatrix}$$

y en las nuevas variables la ecuación puede escribirse como

$$(2-\sqrt{5})x'^2 + (2+\sqrt{5})y'^2 = 6$$

Examinemos la matriz Q. Ya que Q es real y ortogonal, $1 = \det QQ^{-1} = \det QQ^t = \det Q \det Q^t = \det Q \det Q = (\det Q)^2$. De esta manera $\det Q = \pm 1$. Si $\det Q = -1$, podemos intercambiar los renglones de Q para hacer el determinante de esta nueva Q igual a 1. Entonces puede mostrarse (Problema 36) que $Q = \begin{pmatrix} \cos\theta & -\text{sen}\,\theta \\ \text{sen}\,\theta & \cos\theta \end{pmatrix}$ para algún número θ con $0 \le \theta < 2\pi$. Pero, del Ejemplo 5.1.8, esto significa que Q es una matriz de rotación. Por consiguiente, hemos probado el siguiente teorema.

Teorema 1 ***Teorema de los ejes principales en*** $\mathbb{R}^2$. Sea

$$ax^2 + bxy + cy^2 = d \tag{10}$$

una ecuación cuadrática en las variables x y y. Entonces existe un único número θ en $[0, 2\pi)$ tal que la Ecuación (10) puede escribirse en la forma

$$a'x'^2 + c'y'^2 = d \tag{11}$$

donde x', y' son los ejes obtenidos al rotar los ejes x y y un ángulo θ en el sentido contrario a las manecillas del reloj. Más aún, los números a' y c' son los

valores característicos de la matriz $A=\begin{pmatrix} a & b/2 \\ b/2 & c\end{pmatrix}$. Los ejes x' y y' se llaman *ejes principales* de la gráfica de la Ecuación cuadrática (10).

Podemos usar el Teorema 1 para identificar tres secciones cónicas importantes. Recuerde que las *ecuaciones estándares* de la circunferencia, la elipse y la hipérbola son:

Circunferencia:	$x^2+y^2=r^2$	(12)
Elipse:	$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$	(13)
Hipérbola:	$\frac{x^2}{a^2}-\frac{y^2}{b^2}=1$	(14)
	o $\frac{y^2}{a^2}-\frac{x^2}{b^2}=1$	(15)

Ejemplo 2 Identifique la sección cónica cuya ecuación es

$$x^2-4xy+3y^2=6 \tag{16}$$

Solución En el Ejemplo 1 encontramos que esto puede escribirse como $(2-\sqrt{5})x'^2+(2+\sqrt{5})y'^2=6$ o bien

$$\frac{y'^2}{6/(2+\sqrt{5})}-\frac{x'^2}{6/(\sqrt{5}-2)}=1$$

Esta es la Ecuación (15) con $a=\sqrt{6/(2+\sqrt{5})}\approx 1.19$ y $b=\sqrt{6/(\sqrt{5}-2)}\approx 5.04$. Ya que

$$Q=\frac{1}{\sqrt{10-2\sqrt{5}}}\begin{pmatrix} 2 & 1-\sqrt{5} \\ -1+\sqrt{5} & 2\end{pmatrix}$$

y det $Q=1$, tenemos, usando el Problema 36 y el hecho de que 2 y $-1+\sqrt{5}$ son positivos,

$$\cos\theta=\frac{2}{\sqrt{10-2\sqrt{5}}}\approx 0.85065$$

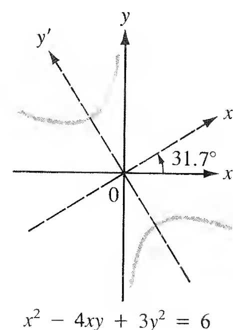


Figura 6.1

De esta manera θ está en el primer cuadrante y, usando una tabla (o una calculadora de bolsillo), encontramos que $\theta\approx 0.5536$ rad $\approx 31.7°$. Por lo tanto, (16) es la ecuación de una hipérbola estándar rotada un ángulo de 31.7° (vea Figura 6.1).

Ejemplo 3 Identifique la sección cónica cuya ecuación es

$$5x^2-2xy+5y^2=4 \tag{17}$$

Solución Aquí $A=\begin{pmatrix} 5 & -1 \\ -1 & 5\end{pmatrix}$, los valores característicos de A son $\lambda_1=4$ y $\lambda_2=6$, y dos vectores característicos ortonormales son $\mathbf{v}_1=\begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2}\end{pmatrix}$ y $\mathbf{v}_2=\begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2}\end{pmatrix}$. Entonces $Q=\begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2}\end{pmatrix}$. Antes de continuar notemos que det $Q=-1$. Para que Q sea una matriz de rotación, necesitamos que det $Q=1$. Esto se consigue fácilmente permutando los vectores característicos. Por consiguiente asignamos $\lambda_1=6$, $\lambda_2=4$, $\mathbf{v}_1=\begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2}\end{pmatrix}$, $\mathbf{v}_2=\begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2}\end{pmatrix}$ y $Q=\begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2}\end{pmatrix}$; ahora bien det $Q=1$. Entonces $D=\begin{pmatrix} 6 & 0 \\ 0 & 4\end{pmatrix}$ y (17) se puede escribir como $D\mathbf{v}\cdot\mathbf{v}=4$ o bien

$$6x'^2+4y'^2=4 \tag{18}$$

donde $\begin{pmatrix} x' \\ y'\end{pmatrix}=Q^t\begin{pmatrix} x \\ y\end{pmatrix}=\begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2}\end{pmatrix}\begin{pmatrix} x \\ y\end{pmatrix}=\begin{pmatrix} 1/\sqrt{2}x-1/\sqrt{2}y \\ 1/\sqrt{2}x+1/\sqrt{2}y\end{pmatrix}$

Simplificando (18), obtenemos $x'^2/(\frac{4}{6})+y'^2/1=1$, lo cual es la Ecuación (13) con $a=\sqrt{\frac{2}{3}}$ y $b=1$. Más aún, ya que $1/\sqrt{2}>0$ y $-1/\sqrt{2}<0$, tenemos, del Problema 36, $\theta=2\pi-\cos^{-1}(1/\sqrt{2})=2\pi-\pi/4=7\pi/4=315°$. De esta manera (17) es la ecuación de una elipse estándar rotada un ángulo de 315° (o 45° en la dirección de las manecillas del reloj). (Vea Figura 6.2.)

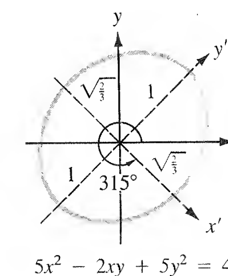


Figura 6.2

Ejemplo 4 Identifique la sección cónica cuya ecuación es

$$-5x^2+2xy-5y^2=4 \tag{19}$$

Solución Recurriendo al Ejemplo 3, la Ecuación (19) puede volver a escribirse como

$$-6x'^2 - 4y'^2 = 4 \tag{20}$$

Puesto que para cualquier número real x' y y' se tiene que $-6x'^2 - 4y'^2 \le 0$, vemos que no hay números reales x y y que satisfagan (19). La sección cónica definida por (19) se llama una *sección cónica degradada* (o *degenerada*).

Existe una manera sencilla de identificar la sección cónica definida por

$$ax^2 + bxy + cy^2 = d \tag{21}$$

Si $A = \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix}$, entonces la ecuación característica de A es

$$\lambda^2 - (a+c)\lambda + (ac - b^2/4) = 0 = (\lambda - \lambda_1)(\lambda - \lambda_2)$$

Esto significa que $\lambda_1\lambda_2 = ac - b^2/4$. Pero la Ecuación (21) puede, como hemos visto, volverse a escribir como

$$\lambda_1 x'^2 + \lambda_2 y'^2 = d \tag{22}$$

Si λ_1 y λ_2 tienen el mismo signo, entonces (21) define una elipse (o una circunferencia) o una cónica degenerada como en los Ejemplos 3 y 4. Si λ_1 y λ_2 tienen signos opuestos, entonces (21) es la ecuación de una hipérbola (Ejemplo 2). Por lo tanto podemos demostrar el siguiente teorema.

Teorema 2 Si $A = \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix}$, entonces la Ecuación cuadrática (21) con $d \neq 0$ es la ecuación de:

i. Una hipérbola si $\det A < 0$.
ii. Una elipse, circunferencia o sección cónica degradada si $\det A > 0$.
iii. Un par de rectas o una sección cónica degradada si $\det A = 0$.
iv. Si $d = 0$, entonces (21) es la ecuación de dos rectas si $\det A \neq 0$, y si $\det A = 0$.

Demostración Ya hemos demostrado que (*i*) y (*ii*) son ciertas. Para probar la parte (*iii*), suponga que $\det A = 0$. Entonces por el Teorema Resumen (Teorema 6.1.6), $\lambda = 0$ es un valor característico de A y la Ecuación (22) se convierte en $\lambda_1 x'^2 = d$ o $\lambda_2 y'^2 = d$. Si $\lambda_1 x'^2 = d$ y $d/\lambda_1 > 0$, entonces $x_1' = \pm\sqrt{d/\lambda_1}$ es la ecuación de dos rectas en el plano xy. Si $d/\lambda_1 < 0$, entonces tenemos $x'^2 < 0$ (lo cual es imposible) y obtenemos una cónica degenerada. Los mismos resultados son válidos si $\lambda_2 y'^2 = d$. La parte (*iv*) se deja como ejercicio (Problema 37). ■

Nota. En el Ejemplo 2 teníamos $\det A = ac - b^24 = -1$. En los Ejemplos 3 y 4 teníamos $\det A = 24$.

Los métodos descritos anteriormente se pueden usar para analizar ecuaciones cuadráticas con más de dos variables. A continuación damos un ejemplo.

Ejemplo 5 Considere la ecuación cuadrática

$$5x^2 + 8xy + 5y^2 + 4xz + 4yz + 2z^2 = 100 \tag{23}$$

Si $A = \begin{pmatrix} 5 & 4 & 2 \\ 4 & 5 & 2 \\ 2 & 2 & 2 \end{pmatrix}$ y $\mathbf{v} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$, entonces (23) puede escribirse en la forma

$$A\mathbf{v} \cdot \mathbf{v} = 100 \tag{24}$$

Del Ejemplo 6.4.2, $Q^tAQ = D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 10 \end{pmatrix}$, donde $Q = \begin{pmatrix} -1/\sqrt{2} & -1/3\sqrt{2} & 2/3 \\ 1/\sqrt{2} & -1/3\sqrt{2} & 2/3 \\ 0 & 4/3\sqrt{2} & 1/3 \end{pmatrix}$

Sea

$$\mathbf{v}' = \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = Q^t\mathbf{v} = \begin{pmatrix} -1/\sqrt{2} & 1/\sqrt{2} & 0 \\ -1/3\sqrt{2} & -1/3\sqrt{2} & 4/3\sqrt{2} \\ 2/3 & 2/3 & 1/3 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

$$= \begin{pmatrix} (-1/\sqrt{2})x + (1/\sqrt{2})y \\ -(1/3\sqrt{2})x - (1/3\sqrt{2})y + (4/3\sqrt{2})z \\ (2/3)x + (2/3)y + (1/3)z \end{pmatrix}$$

Entonces, como antes $A = QDQ^t$ y $A\mathbf{v} \cdot \mathbf{v} = QDQ^t\mathbf{v} \cdot \mathbf{v} = DQ^t\mathbf{v} \cdot Q^t\mathbf{v} = D\mathbf{v}' \cdot \mathbf{v}'$. De esta manera (24) puede escribirse con las nuevas variables x', y, z' como $D\mathbf{v}' \cdot \mathbf{v}' = 100$ o bien

$$x'^2 + y'^2 + 10z'^2 = 100 \tag{25}$$

En $\mathbb{R}^3$, la superficie definida por (25) se llama *elipsoide* (Figura 6.3).

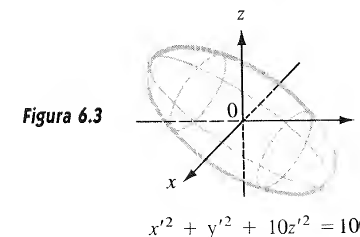


Figura 6.3 $x'^2 + y'^2 + 10z'^2 = 100$

Existe una amplia variedad de superficies de tres dimensiones con la forma $A\mathbf{v} \cdot \mathbf{v} = d$, donde $\mathbf{v} \in \mathbb{R}^2$. Estas superficies se llaman *superficies cuádricas*.

Cerraremos esta sección haciendo notar que las formas cuadráticas pueden estar definidas en cualquier número de variables.

Definición 2 **Forma cuadrática.** Sea $\mathbf{v}=\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$, y sea A una matriz simétrica de $n \times n$. Entonces una *forma cuadrática* en $x_1, x_2, \ldots, x_n$ es una expresión de la forma

$$F(x_1, x_2, \ldots, x_n) = A\mathbf{v}\cdot\mathbf{v} \tag{26}$$

Ejemplo 6 Sea

$$A=\begin{pmatrix} 2 & 1 & 2 & -2 \\ 1 & -4 & 6 & 5 \\ 2 & 6 & 7 & -1 \\ -2 & 5 & -1 & 3 \end{pmatrix}$$

Entonces
$$A\mathbf{v}\cdot\mathbf{v}=\left[\begin{pmatrix} 2 & 1 & 2 & -2 \\ 1 & -4 & 6 & 5 \\ 2 & 6 & 7 & -1 \\ -2 & 5 & -1 & 3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}\right]\cdot\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$$

$$=\begin{pmatrix} 2x_1+x_2+2x_3-2x_4 \\ x_1-4x_2+6x_3+5x_4 \\ 2x_1+6x_2+7x_3-x_4 \\ -2x_1+5x_2-x_3+3x_4 \end{pmatrix}\cdot\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$$

$$=2x_1^2+2x_1x_2-4x_2^2+4x_1x_3+12x_2x_3$$
$$+7x_3^2-4x_1x_4+10x_2x_4-2x_3x_4+3x_4^2$$

(después de simplificar)

Ejemplo 7 Encuentre la matriz simétrica A correspondiente a la forma cuadrática

$$5x_1^2-3x_1x_2+4x_2^2+8x_1x_3-9x_2x_3+2x_3^2-x_1x_4+7x_2x_4+6x_3x_4+9x_4^2$$

Solución Si $A=(a_{ij})$, entonces observando los ejemplos anteriores de esta sección, vemos que a_{ii} es el coeficiente del término x_i^2 y $a_{ij}+a_{ji}$ es el coeficiente del término x_ix_j. Puesto que A es simétrica, $a_{ij}=a_{ji}$; por lo tanto $a_{ij}=a_{ji}=\frac{1}{2}\cdot$ (coeficiente del término x_ix_j). Escribiendo todo esto junto, obtenemos

$$A=\begin{pmatrix} 5 & -\frac{3}{2} & 4 & -\frac{1}{2} \\ -\frac{3}{2} & 4 & -\frac{9}{2} & \frac{7}{2} \\ 4 & -\frac{9}{2} & 2 & 3 \\ -\frac{1}{2} & \frac{7}{2} & 3 & 9 \end{pmatrix}$$

Problemas 6.5

En los Problemas del 1 al 13 escriba la ecuación cuadrática en la forma $A\mathbf{v}\cdot\mathbf{v}=d$ (donde A es una matriz simétrica) y elimine el término xy girando los ejes un ángulo θ. Escriba la ecuación en términos de las nuevas variables e identifique la sección cónica obtenida.

1. $3x^2-2xy-5=0$
2. $4x^2+4xy+y^2=9$
3. $4x^2+4xy-y^2=9$
4. $xy=1$
5. $xy=a;\ a>0$
6. $4x^2+2xy+3y^2+2=0$
7. $xy=a;\ a<0$
8. $x^2+4xy+4y^2-6=0$
9. $-x^2+2xy-y^2=0$
10. $2x^2+xy+y^2=4$
11. $3x^2-6xy+5y^2=36$
12. $x^2-3xy+4y^2=1$
13. $6x^2+5xy-6y^2+7=0$
14. ¿Cuáles son las posibles formas de la gráfica de $ax^2+bxy+cy^2=0$?

En los Problemas del 15 al 18 escriba la forma cuadrática en las nuevas variables x', y' y z' de modo que no aparezcan términos cruzados (xy, xz, yz).

15. $x^2-2xy+y^2-2xz-2yz+z^2$
16. $-x^2+4xy-y^2+4xz+4yz+z^2$
17. $3x^2+4xy+2y^2+4xz+4z^2$
18. $x^2-2xy+2y^2-2yz+z^2$

En los Problemas del 19 al 21 encuentre una matriz simétrica A tal que la forma cuadrática pueda ser escrita en la forma $A\mathbf{x}\cdot\mathbf{x}$.

19. $x_1^2+2x_1x_2+x_2^2+4x_1x_3+6x_2x_3+3x_3^2+7x_1x_4-2x_2x_4+x_4^2$
20. $x_1^2-x_2^2+x_1x_3-x_2x_4+x_4^2$
21. $3x_1^2-7x_1x_2-2x_2^2+x_1x_3-x_2x_3+3x_3^2-2x_1x_4+x_2x_4-4x_3x_4-6x_4^2+3x_1x_5-5x_3x_5+x_4x_5-x_5^2$
22. Suponga que para algún valor de d distinto de cero, la gráfica de $ax^2+bxy+cy^2=d$ es una hipérbola. Muestre que la gráfica es una hipérbola para cualquier otro valor de d distinto de cero.
23. Muestre que si $a\neq c$, el término xy en la ecuación cuadrática (1) será eliminado por una rotación de un ángulo θ, en donde θ está dado por $\cot 2\theta=(a-c)/b$.
24. Muestre que si $a=c$ en el Problema 23, entonces el término xy será eliminado por una rotación de un ángulo de $\pi/4$ o bien $-\pi/4$.
25. ★ Suponga que una rotación convierte $ax^2+bxy+cy^2$ en $a'(x')^2+b'(x'y')+c'(y')^2$. Muestre que:
26. Se dice que una forma cuadrática $F(\mathbf{x})=F(x_1, x_2, \ldots, x_n)$ es *definida positiva* si $F(\mathbf{x})\geq 0$ para todo $\mathbf{x}\in\mathbb{R}^n$ y $F(\mathbf{x})=0$ si y sólo si $\mathbf{x}=0$. Muestre que F es definida positiva si y sólo si la matriz simétrica A asociada con F tiene valores característicos positivos.
27. Se dice que una forma cuadrática $F(\mathbf{x})$ es *semidefinida positiva* si $F(\mathbf{x})\geq 0$ para todo $\mathbf{x}\in\mathbb{R}^n$. Muestre que F es semidefinida positiva si y sólo si los valores característicos de la matriz simétrica asociada con F son todos no negativos.

Las definiciones de definida negativa, semidefinida negativa *son las de los Problemas 26 y 27 con ≤ 0 en lugar de ≥ 0. Una forma cuadrática es* indefinida *si no es ninguna de las anteriores. En los Problemas del 28 al 35 determine si la forma cuadrática dada es definida positiva, semidefinida positiva, definida negativa o indefinida.*

28. $3x^2+2y^2$
29. $-3x^2-3y^2$
30. $3x^2-2y^2$
31. $x^2+2xy+2y^2$

32. $x^2-2xy+2y^2$ **33.** $x^2-4xy+3y^2$ **34.** $-x^2+4xy-3y^2$ **35.** $-x^2+2xy-2y^2$

★ **36.** Sea $Q=\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ una matriz ortogonal real con det Q = 1. Defina el número $\theta \in [0, 2\pi)$:

a. Si $a \geq 0$ y $c > 0$, entonces $\theta = \cos^{-1} a$ $(0 < \theta \leq \pi/2)$.
b. Si $a \geq 0$ y $c < 0$, entonces $\theta = 2\pi - \cos^{-1} a$ $(3\pi/2 \leq \theta < 2\pi)$.
c. Si $a \leq 0$ y $c > 0$, entonces $\theta = \cos^{-1} a$ $(\pi/2 \leq \theta < \pi)$.
d. Si $a \leq 0$ y $c < 0$, entonces $\theta = 2\pi - \cos^{-1} a$ $(\pi < \theta \leq 3\pi/2)$.
e. Si $a = 1$ y $c = 0$, entonces $\theta = 0$.
f. Si $a = -1$ y $c = 0$, entonces $\theta = \pi$.

(Aquí $\cos^{-1} x \in [0, \pi]$ para $x \in [-1, 1]$.) Con θ escogido como antes muestre que

$$Q=\begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}.$$

37. Demuestre, usando la fórmula 22, que la Ecuación (21) es la ecuación de dos rectas en el plano xy cuando $d = 0$ y det $A \neq 0$. Si det $A = d = 0$ muestre que la Ecuación (21) es la ecuación de una sola recta.

38. Sea A la representación matricial simétrica de la ecuación cuadrática (1) con $d \neq 0$. Sean λ_1 y λ_2 los valores característicos de A. Muestre que (1) es la ecuación de: **(a)** Una hipérbola si $\lambda_1\lambda_2 < 0$ y **(b)** una circunferencia, elipse o sección cónica degenerada si $\lambda_1\lambda_2 > 0$.

6.6 Forma canónica de Jordan

Como hemos visto, las matrices de $n \times n$ con n vectores característicos linealmente independientes pueden ser llevados a una forma especialmente agradable mediante una transformación de equivalencia. Afortunadamente, como la "mayoría" de los polinomios tienen diferentes raíces, la "mayoría" de las matrices tendrán diferentes valores característicos. Sin embargo, como veremos en la Sección 6.7, las matrices que no son diagonalizables (esto es, que no tienen n vectores característicos linealmente independientes) aparecen en ciertas aplicaciones. En este caso todavía es posible mostrar que la matriz es equivalente a otra matriz más simple, pero la nueva matriz ya no es diagonal y la matriz de transformación C es más difícil de obtener.

Para discutir este caso completamente, definimos la matriz N_k como la matriz de $k \times k$

$$N_k=\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix} \qquad (1)$$

Note que N_k es la matriz con unos arriba de la diagonal principal y ceros en las

demás posiciones. Definimos una *matriz bloque de Jordan** $B(\lambda)$ de $k \times k$ como

$$B(\lambda)=\lambda I+N_k=\begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & & \cdots & \lambda & 1 \\ 0 & 0 & & \cdots & 0 & \lambda \end{pmatrix} \qquad (2)$$

Esto es, $B(\lambda)$ es la matriz de $k \times k$ con el número fijo λ en la diagonal, unos arriba de la diagonal y ceros en las demás posiciones.

Nota. Podemos tener (y con frecuencia tendremos) una matriz bloque de Jordan de 1×1, que toma la forma $B(\lambda) = (\lambda)$.

Una *matriz de Jordan* J tiene la forma

$$J=\begin{pmatrix} B_1(\lambda_1) & 0 & \cdots & 0 \\ 0 & B_2(\lambda_2) & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & B_r(\lambda_r) \end{pmatrix}$$

donde cada $B_j(\lambda_j)$ es una matriz bloque de Jordan. De esta manera *una matriz de Jordan es una matriz con matrices bloque de Jordan en la diagonal y ceros en las demás componentes.*

Ejemplo 1 Los siguientes ejemplos muestran matrices de Jordan. Los bloques de Jordan están marcados por las líneas punteadas

i. $\left(\begin{array}{cc:c} 2 & 1 & 0 \\ 0 & 2 & 0 \\ \hdashline 0 & 0 & 4 \end{array}\right)$

ii. $\left(\begin{array}{c:ccc:c} -3 & 0 & 0 & 0 & 0 \\ \hdashline 0 & -3 & 1 & 0 & 0 \\ 0 & 0 & -3 & 1 & 0 \\ 0 & 0 & 0 & -3 & 0 \\ \hdashline 0 & 0 & 0 & 0 & 7 \end{array}\right)$

* Llamada así en honor del matemático francés Camille Jordan (1838-1922). Los conceptos de esta sección aparecieron por primera vez en su excelente trabajo *Traité des substitutions et des équations algebriques*, el cual fue publicado en 1870.

iii. $\begin{pmatrix} 4 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 5 \end{pmatrix}$

Ejemplo 2 Las únicas matrices de Jordan de 2×2 son $\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$ y $\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$.

Ejemplo 3 Las únicas matrices de Jordan de 3×3 son

$$\begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \quad \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 1 \\ 0 & 0 & \lambda_2 \end{pmatrix} \quad \begin{pmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{pmatrix} \quad \begin{pmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 1 \\ 0 & 0 & \lambda_1 \end{pmatrix}$$

donde λ_1, λ_2 y λ_3 no son necesariamente distintas.

El siguiente resultado es uno de los teoremas más importantes en la teoría de matrices. Aunque su demostración está más allá del alcance de este libro*, probaremos este teorema en el caso de 2×2 (vea Teorema 3) y sugeriremos una demostración para el caso de 3×3 en el Problema 19.

Teorema 1 Sea A una matriz de $n\times n$. Entonces existe una matriz invertible C de $n\times n$ tal que

$$C^{-1}AC = J \tag{3}$$

donde J es una matriz de Jordan cuyos elementos diagonales son los valores característicos de A. Más aún, J es única excepto por el orden en que aparecen los bloques de Jordan.

Observación 1. La última afirmación del teorema significa que si A es equivalente a

* Si desea ver una demostración consulte G. Birkhoff y S. MacLane, *A Survey of Modern Algebra* (Nueva York: MacMillan, 1953), p. 334.

$J_1 = \begin{pmatrix} 2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 1 & 0 & 0 \\ 0 & 0 & 0 & 3 & 1 & 0 \\ 0 & 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 4 \end{pmatrix}$, entonces A también es equivalente a

$$J_2 = \begin{pmatrix} 3 & 1 & 0 & 0 & 0 & 0 \\ 0 & 3 & 1 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 2 \end{pmatrix} \quad \text{y} \quad J_3 = \begin{pmatrix} 4 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 1 & 0 \\ 0 & 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 0 & 3 \end{pmatrix}$$

y otras tres matrices de Jordan. Esto es, los bloques de Jordan son los mismos, pero se puede cambiar el orden en el que se escriben.

Definición 1 La matriz J se llama *forma canónica de Jordan de A*.

Observación. Si A es diagonalizable, entonces, $J = D = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$, donde $\lambda_1, \lambda_2, \ldots, \lambda_n$ son los valores característicos de A (no necesariamente distintos). Cada componente de la diagonal es una matriz bloque de Jordan de 1×1.

Veremos ahora cómo calcular la forma canónica de Jordan de cualquier matriz de 2×2. Si A tiene dos vectores característicos linealmente independientes, entonces ya sabemos qué hacer. Por consiguiente el único caso de interés ocurre cuando A tiene un solo valor característico λ de multiplicidad algebraica 2 y multiplicidad geométrica 1. Esto es, suponemos que, para λ, A tiene el único vector característico independiente $\mathbf{v}_1$. Esto es: *Cualquier vector que no sea un múltiplo de* $\mathbf{v}_1$ *no es un vector característico.*

Teorema 2 Sea la matriz A de 2×2 con un valor característico λ de multiplicidad algebraica 2 y multiplicidad geométrica 1. Sea $\mathbf{v}_1$ un vector característico correspondiente a λ. Entonces existe un vector $\mathbf{v}_2$ que satisface la ecuación

$$(A - \lambda I)\mathbf{v}_2 = \mathbf{v}_1 \tag{4}$$

Demostración Sea $\mathbf{x} \in \mathbb{C}^2$ un vector constante que no sea un múltiplo de $\mathbf{v}_1$ por lo cual $\mathbf{x}$ no es un vector característico de A. Primero mostraremos que

$$\mathbf{w} = (A - \lambda I)\mathbf{x} \tag{5}$$

es un vector característico de A. Esto es, demostraremos que $\mathbf{w} = c\mathbf{v}_1$ para alguna constante c. Ya que $\mathbf{w} \in \mathbb{C}^2$ y $\mathbf{v}_1$ y $\mathbf{x}$ son linealmente independientes, existen constantes c_1 y c_2 tales que

$$\mathbf{w} = c_1\mathbf{v}_1 + c_2\mathbf{x} \tag{6}$$

Para mostrar que $\mathbf{w}$ es un vector característico de A, debemos mostrar que $c_2 = 0$. De (5) y (6) encontramos que

$$(A - \lambda I)\mathbf{x} = c_1\mathbf{v}_1 + c_2\mathbf{x} \tag{7}$$

Sea $B = A - (\lambda + c_2)I$. Entonces, de (7),

$$B\mathbf{x} = [A - (\lambda + c_2)I]\mathbf{x} = c_1\mathbf{v}_1 \tag{8}$$

Si suponemos que $c_2 \neq 0$ entonces $\lambda + c_2 \neq \lambda$ y $\lambda + c_2$ no es un valor característico de A (ya que λ es el único valor característico de A). Se tiene de esta forma que $\det B = \det[A - (\lambda + c_2)I] \neq 0$, lo cual significa que B es invertible. De aquí que (8) puede escribirse como

$$\mathbf{x} = B^{-1}c_1\mathbf{v}_1 = c_1B^{-1}\mathbf{v}_1 \tag{9}$$

Entonces, multiplicando ambos lados de (9) por λ,

$$\lambda\mathbf{x} = \lambda c_1B^{-1}\mathbf{v}_1 = c_1B^{-1}\lambda\mathbf{v}_1 = c_1B^{-1}A\mathbf{v}_1 \tag{10}$$

Pero $B = A - (\lambda + c_2)I$, de manera que

$$A = B + (\lambda + c_2)I \tag{11}$$

Sustituyendo (11) en (10),

$$\begin{aligned} \lambda\mathbf{x} &= c_1B^{-1}[B + (\lambda + c_2)I]\mathbf{v}_1 \\ &= c_1[I + (\lambda + c_2)B^{-1}]\mathbf{v}_1 \\ &= c_1\mathbf{v}_1 + (\lambda + c_2)c_1B^{-1}\mathbf{v}_1 \end{aligned} \tag{12}$$

Pero, usando de nuevo (8), $c_1B^{-1}\mathbf{v}_1 = \mathbf{x}$, por lo que (12) se convierte en

$$\lambda\mathbf{x} = c_1\mathbf{v}_1 + (\lambda + c_2)\mathbf{x} = c_1\mathbf{v}_1 + c_2\mathbf{x} + \lambda\mathbf{x}$$

o bien

$$\mathbf{0} = c_1\mathbf{v}_1 + c_2\mathbf{x} \tag{13}$$

Pero $\mathbf{v}_1$ y $\mathbf{x}$ son linealmente independientes, así $c_1 = c_2 = 0$. Esto contradice la hipótesis de que $c_2 \neq 0$. De esta manera $c_2 = 0$ y, por (6), $\mathbf{w}$ es un múltiplo de $\mathbf{v}_1$ por lo que $\mathbf{w} = c_1\mathbf{v}_1$ es un vector característico de A. Más aún, $\mathbf{w} \neq \mathbf{0}$ ya que si $\mathbf{w} = \mathbf{0}$ entonces (5) nos dice que $\mathbf{x}$ es un vector característico de A. Por consiguiente $c_1 \neq 0$. Sea

$$\mathbf{v}_2 = \frac{1}{c_1}\mathbf{x} \tag{14}$$

Entonces $(A - \lambda I)\mathbf{v}_2 = (1/c_1)(A - \lambda I)\mathbf{x} = (1/c_1)\mathbf{w} = \mathbf{v}_1$. Esto demuestra el teorema. ■

Definición 2 ***Vector característico generalizado.*** Sea A una matriz de 2×2 con el único valor característico λ con multiplicidad geométrica 1. Sea $\mathbf{v}_1$ un valor característico de A. Entonces el vector $\mathbf{v}_2$ definido por $(A - \lambda I)\mathbf{v}_2 = \mathbf{v}_1$ se llama *vector característico generalizado* de A correspondiente al valor característico de λ.

Ejemplo 4 Sea $A = \begin{pmatrix} 3 & -2 \\ 8 & -5 \end{pmatrix}$. La ecuación característica de A es $\lambda^2 + 2\lambda + 1 = (\lambda + 1)^2 = 0$, por lo que $\lambda = -1$ es un valor característico de multiplicidad algebraica 2. Entonces

$$(A - \lambda I)\mathbf{v} = (A + I)\mathbf{v} = \begin{pmatrix} 4 & -2 \\ 8 & -4 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

De aquí obtenemos el vector característico $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$. No existe otro vector característico linealmente independiente. Para encontrar un vector característico generalizado $\mathbf{v}_2$, calculamos $(A + I)\mathbf{v}_2 = \mathbf{v}_1$ o bien $\begin{pmatrix} 4 & -2 \\ 8 & -4 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, lo cual nos da el sistema

$$\begin{aligned} 4x_1 - 2x_2 &= 1 \\ 8x_1 - 4x_2 &= 2 \end{aligned}$$

La segunda ecuación es el doble de la primera, por lo que x_2 puede escogerse arbitrariamente y $x_1 = (1 + 2x_2)/4$. Por consiguiente una posible elección de $\mathbf{v}_2$ es $\mathbf{v}_2 = \begin{pmatrix} \frac{1}{4} \\ 0 \end{pmatrix}$.

La necesidad de encontrar el vector característico generalizado se da en el siguiente teorema.

Teorema 3 Sea A, λ, $\mathbf{v}_1$ y $\mathbf{v}_2$ como en el Teorema 2 y sea C la matriz cuyas columnas son $\mathbf{v}_1$ y $\mathbf{v}_2$. Entonces $C^{-1}AC = J$, en donde $J = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$ es la forma canónica de Jordan de A.

Demostración Puesto que $\mathbf{v}_1$ y $\mathbf{v}_2$ son linealmente independientes, vemos que C es invertible. Ahora nótese que $AC = A(\mathbf{v}_1, \mathbf{v}_2) = (A\mathbf{v}_1, A\mathbf{v}_2) = (\lambda\mathbf{v}_1, A\mathbf{v}_2)$. Pero por la Ecuación (4), $A\mathbf{v}_2 = \mathbf{v}_1 + \lambda\mathbf{v}_2$ de modo que $AC = (\lambda\mathbf{v}_1, \mathbf{v}_1 + \lambda\mathbf{v}_2)$. Asimismo $CJ = (\mathbf{v}_1, \mathbf{v}_2)\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} = (\lambda\mathbf{v}_1, \mathbf{v}_1 + \lambda\mathbf{v}_2)$. De esta manera $AC = CJ$, lo cual significa que $C^{-1}AC = J$ y el teorema está demostrado. ■

Ejemplo 5 En el Ejemplo 4, $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ y $\mathbf{v}_2 = \begin{pmatrix} \frac{1}{4} \\ 0 \end{pmatrix}$. Entonces $C = \begin{pmatrix} 1 & \frac{1}{4} \\ 2 & 0 \end{pmatrix}$, $C^{-1} = -2\begin{pmatrix} 0 & -\frac{1}{4} \\ -2 & 1 \end{pmatrix} = \begin{pmatrix} 0 & \frac{1}{2} \\ 4 & -2 \end{pmatrix}$, y

$$C^{-1}AC = \begin{pmatrix} 0 & \frac{1}{2} \\ 4 & -2 \end{pmatrix}\begin{pmatrix} 3 & -2 \\ 8 & -5 \end{pmatrix}\begin{pmatrix} 1 & \frac{1}{4} \\ 2 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & \frac{1}{2} \\ 4 & -2 \end{pmatrix}\begin{pmatrix} -1 & \frac{3}{4} \\ -2 & 2 \end{pmatrix} = \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix} = J$$

El método antes descrito puede generalizarse para obtener la forma canónica de Jordan de cada matriz. Nosotros no haremos esto, pero sugerimos una generalización en el Problema 19. Aunque no demostraremos este resultado siempre es posible determinar el número de unos arriba de la diagonal en la forma canónica de Jordan de una matriz de A de $n \times n$. Sea λ_i un valor característico de A con multiplicidad algebraica r_i y multiplicidad geométrica s_i. Si λ_1, $\lambda_2, \ldots, \lambda_k$ son los valores característicos de A, entonces:

> El número de unos arriba de la diagonal de la forma canónica de Jordan de A
>
> $$= (r_1 - s_1) + (r_2 - s_2) + \cdots + (r_k - s_k) \qquad (15)$$
> $$= \sum_{i=1}^{k} r_i - \sum_{i=1}^{k} s_i = n - \sum_{i=1}^{k} s_i$$

Si conocemos la ecuación característica de una matriz A, entonces podemos determinar las posibles formas canónicas de Jordan de A.

Ejemplo 6 Si la ecuación característica de A es $(\lambda-2)^3(\lambda+3)$, entonces las posibles formas canónicas de Jordan de A son

$$J = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & -3 \end{pmatrix}, \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & -3 \end{pmatrix}, \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & -3 \end{pmatrix}$$

o cualquier matriz obtenida por un nuevo arreglo de los bloques de Jordan en J. La primera matriz corresponde a una multiplicidad geométrica de 3 (para $\lambda=2$); la segunda corresponde a una multiplicidad geométrica de 2; y la tercera corresponde a una multiplicidad geométrica de 1.

Problemas 6.6

En los Problemas del 1 *al* 14 *determine si la matriz dada es una matriz de Jordan.*

1. $\begin{pmatrix} 1 & 1 \\ 0 & -6 \end{pmatrix}$ 2. $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ 3. $\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$ 4. $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{pmatrix}$

5. $\begin{pmatrix} 3 & 1 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{pmatrix}$ 6. $\begin{pmatrix} 3 & 1 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 2 \end{pmatrix}$ 7. $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 4 \end{pmatrix}$ 8. $\begin{pmatrix} 1 & 1 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{pmatrix}$

9. $\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$ 10. $\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix}$ 11. $\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$

12. $\begin{pmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & 3 & 1 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 0 & 5 \end{pmatrix}$ 13. $\begin{pmatrix} a & 0 & 0 & 0 & 0 \\ 0 & b & 0 & 0 & 0 \\ 0 & 0 & c & 0 & 0 \\ 0 & 0 & 0 & d & 0 \\ 0 & 0 & 0 & 0 & e \end{pmatrix}$ 14. $\begin{pmatrix} a & 1 & 0 & 0 & 0 \\ 0 & a & 0 & 0 & 0 \\ 0 & 0 & c & 1 & 0 \\ 0 & 0 & 0 & c & 1 \\ 0 & 0 & 0 & 0 & c \end{pmatrix}$

En los Problemas del 15 *al* 18 *encuentre una matriz invertible C que transforme la matriz de* 2 × 2 *a su forma canónica de Jordan.*

15. $\begin{pmatrix} 6 & 1 \\ 0 & 6 \end{pmatrix}$ 16. $\begin{pmatrix} -12 & 7 \\ -7 & 2 \end{pmatrix}$ 17. $\begin{pmatrix} -10 & -7 \\ 7 & 4 \end{pmatrix}$ 18. $\begin{pmatrix} 4 & -1 \\ 1 & 2 \end{pmatrix}$

★ **19.** Sea A una matriz de 3 × 3. Suponga que λ es un valor característico de A con multiplicidad algebraica 3 y multiplicidad geométrica 1 y sea $\mathbf{v}_1$ el vector característico correspondiente.
   a. Muestre que existe una solución, $\mathbf{v}_2$, para el sistema $(A - \lambda I)\mathbf{v}_2 = \mathbf{v}_1$ tal que $\mathbf{v}_1$ y $\mathbf{v}_2$ son linealmente independientes.
   b. Con $\mathbf{v}_2$ definida por la parte (a), demuestre que existe una solución $\mathbf{v}_3$ para el sistema $(A - \lambda I)\mathbf{v}_3 = \mathbf{v}_2$ tal que $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$ son linealmente independientes.
   c. Demuestre que si C es una matriz cuyas columnas son $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$, entonces

$$C^{-1}AC = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}.$$

20. Aplique el procedimiento descrito en el Problema 19 para reducir la matriz A = $A = \begin{pmatrix} -2 & 1 & 0 \\ -2 & 1 & -1 \\ -1 & 1 & -2 \end{pmatrix}$ mediante una transformación equivalente a su forma canónica de Jordan.

21. Haga lo mismo para $A = \begin{pmatrix} -1 & -2 & -1 \\ -1 & -1 & -1 \\ 2 & 3 & 2 \end{pmatrix}$.

22. Haga lo mismo para $A = \begin{pmatrix} -1 & -18 & -7 \\ 1 & -13 & -4 \\ -1 & 25 & 8 \end{pmatrix}$.

23. Una matriz A de $n \times n$ es nulipotente si existe un entero k tal que $A^k = 0$. Si k es el entero más pequeño con tal característica, entonces k es llamado el *índice de nulipotencia* de A. Demuestre que si k es el índice de nulipotencia de A y si $m \geq k$ entonces $A^m = 0$.

24. Sea N_k la matriz definida por la Ecuación (1). Demuestre que N_k es nulipotente con índice de nulipotencia k.

25. Escriba todas las posibles matrices de Jordan de 4 × 4.

En los Problemas del 26 *al* 31 *está dado el polinomio característico de una matriz A. Escriba las posibles formas canónicas de Jordan para A.*

26. $(\lambda+1)^2(\lambda-2)^2$ **27.** $(\lambda-3)^3(\lambda+4)$ **28.** $(\lambda-3)^4$
29. $(\lambda-4)^3(\lambda+3)^2$ **30.** $(\lambda-6)(\lambda+7)^4$ **31.** $(\lambda+7)^5$

32. Usando la forma canónica de Jordan, demuestre que, para cualquier matriz A de $n \times n$, $\det A = \lambda_1\lambda_2 \cdots \lambda_n$, donde $\lambda_1, \lambda_2, \ldots, \lambda_n$ son los valores característicos de A.

6.7 Una aplicación importante: Ecuaciones diferenciales matriciales

Sea $x=f(t)$ una función que representa una cantidad física tal como el volumen de una sustancia, la población de determinada especie, la masa de una sustancia radiactiva o el número de pesos invertidos en bonos. Entonces el crecimiento de $f(t)$ está dado por su derivada $f'(t)=dx/dt$. Si $f(t)$ crece a una razón constante, entonces $dx/dt=k$ y $x=kt+C$; esto es $x=f(t)$ es una recta.

Con frecuencia es más interesante y más apropiado el considerar la *tasa de crecimiento relativo* definida por

$$\text{Tasa de crecimiento relativo} = \frac{\text{tamaño real del crecimiento}}{\text{tamaño de } f(t)} = \frac{f'(t)}{f(t)} = \frac{x'(t)}{x(t)} \qquad (1)$$

Si la tasa de crecimiento relativo es constante, entonces

$$\frac{x'(t)}{x(t)} = a \qquad (2)$$

o bien

$$x'(t) = ax(t) \qquad (3)$$

La Ecuación (3) se llama *ecuación diferencial* ya que es una ecuación que tiene una derivada. No es difícil demostrar que las únicas soluciones de (3) son de la forma

$$x(t) = ce^{at} \qquad (4)$$

donde c es una constante arbitraria. Sin embargo, si $x(t)$ representa una cantidad física, entonces normalmente se especifica un *valor inicial* $x_0 = x(0)$ de la cantidad. Entonces, sustituyendo $t = 0$ en (4), tenemos $x_0 = x(0) = ce^{a\cdot 0} = c$ o bien

$$x(t) = x_0e^{at} \qquad (5)$$

La función $x(t)$ dada por (5) es la única solución de (3) que satisface la condición inicial $x(0)=x_0$.

La Ecuación (3) surge en muchas aplicaciones interesantes. Algunas de ellas aparecen indudablemente en los textos de Cálculo, en el capítulo relativo a la

función exponencial. En esta sección consideraremos una generalización de la Ecuación (3).

En el modelo discutido anteriormente, buscamos una función desconocida. A menudo se tiene que aparecen varias funciones vinculadas por varias ecuaciones diferenciales. Algunos ejemplos se dan más adelante en esta sección. Considere el siguiente sistema de n ecuaciones diferenciales en n funciones incógnitas

$$\begin{aligned} x_1'(t) &= a_{11}x_1(t)+a_{12}x_2(t)+\cdots+a_{1n}x_n(t) \\ x_2'(t) &= a_{21}x_1(t)+a_{22}x_2(t)+\cdots+a_{2n}x_n(t) \\ &\vdots \\ x_n'(t) &= a_{n1}x_1(t)+a_{n2}x_2(t)+\cdots+a_{nn}x_n(t) \end{aligned} \qquad (6)$$

donde los a_{ij} son números reales. El sistema (6) se denomina un *sistema de ecuaciones diferenciales lineales de primer orden de* $n\times n$. El término "primer orden" significa que en el sistema únicamente aparecen las primeras derivadas.

Ahora, sea

$$\mathbf{x}(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{pmatrix}$$

Aquí $\mathbf{x}(t)$ se llama *función vectorial*. Definimos

$$\mathbf{x}'(t) = \begin{pmatrix} x_1'(t) \\ x_2'(t) \\ \vdots \\ x_n'(t) \end{pmatrix}$$

Entonces, si definimos la matriz de $n\times n$

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

el sistema (6) puede escribirse como

$$\mathbf{x}'(t) = A\mathbf{x}(t) \qquad (7)$$

Note que la Ecuación (7) es casi idéntica a la Ecuación (3). La única diferencia es que ahora tenemos una función vectorial y una matriz, mientras que antes teníamos una función "escalar" y un número (matriz de 1×1).

Para resolver la Ecuación (7) debemos conjeturar que una solución tendría la

forma e^{At}. Pero ¿qué significa e^{At}? Responderemos a esta pregunta en un momento. Primero, recordemos el desarrollo en serie de la función e^t:

$$e^t = 1 + t + \frac{t^2}{2!} + \frac{t^3}{3!} + \frac{t^4}{4!} + \cdots \tag{8}$$

Esta serie converge para todo número real t. Entonces, para todo número real a,

$$e^{at} = 1 + at + \frac{(at)^2}{2!} + \frac{(at)^3}{3!} + \frac{(at)^4}{4!} + \cdots \tag{9}$$

Definición 1 **La matriz e^A.** Sea A una matriz de $n \times n$ con componentes reales (o complejas). Entonces e^A es una matriz de $n \times n$ definida por

$$e^A = I + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \frac{A^4}{4!} + \cdots \tag{10}$$

Observación. No es difícil demostrar que la serie de matrices en la Ecuación (10) converge para cada matriz A, pero hacerlo nos desviaría de nuestro objetivo. Sin embargo, podemos dar una indicación de por qué sucede esto. Primero definimos $|A|_i$ como la suma de los valores absolutos de las componentes en el renglón i-ésimo de A. Entonces definimos la *norma* de A, denotada $|A|$, por

$$|A| = \max_{1 \le i \le n} |A|_i \tag{11}$$

Se puede mostrar que

$$|AB| \le |A|\,|B| \tag{12}$$

y

$$|A + B| \le |A| + |B| \tag{13}$$

Entonces, usando (12) y (13) en (10), obtenemos

$$|e^A| \le 1 + |A| + \frac{|A|^2}{2!} + \frac{|A|^3}{3!} + \frac{|A|^4}{4!} + \cdots = e^{|A|}$$

Puesto que $|A|$ es un número real, $e^{|A|}$ es finito. Esto demuestra que la serie en (10) converge para toda matriz A.

Ahora veremos la utilidad de la serie en la Ecuación (10).

Teorema 1 Para cualquier vector constante $\mathbf{c}$, $\mathbf{x}(t) = e^{At}\mathbf{c}$ es una solución de (7). Además, la solución de (7) dada por $\mathbf{x}(t) = e^{At}\mathbf{x}_0$ satisface $\mathbf{x}(0) = \mathbf{x}_0$.

Demostración Calculamos, usando (10):

$$\mathbf{x}(t) = e^{At}\mathbf{c} = \left[I + At + A^2\frac{t^2}{2!} + A^3\frac{t^3}{3!} + \cdots\right]\mathbf{c} \tag{14}$$

Pero como A es una matriz constante y $\mathbf{c}$ es un vector constante tenemos

$$\frac{d}{dt}A^k\frac{t^k}{k!} = \frac{d}{dt}\frac{t^k}{k!}A^k = \frac{kt^{k-1}}{k!}A^k = \frac{A^k t^{k-1}}{(k-1)!} = A\left[A^{k-1}\frac{t^{k-1}}{(k-1)!}\right] \tag{15}$$

Entonces, combinando (14) y (15), obtenemos (debido a que $\mathbf{c}$ es un vector constante)

$$\mathbf{x}'(t) = \frac{d}{dt}e^{At}\mathbf{c} = A\left[I + At + A^2\frac{t^2}{2!} + A^3\frac{t^3}{3!} + \cdots\right]\mathbf{c} = Ae^{At}\mathbf{c} = A\mathbf{x}(t)$$

Finalmente, en virtud de que $e^{A\cdot 0} = e^0 = I$,

$$\mathbf{x}(0) = e^{A\cdot 0}\mathbf{x}_0 = I\mathbf{x}_0 = \mathbf{x}_0. \blacksquare$$

Definición 2 **Matriz solución principal.** La matriz e^{At} se llama *matriz solución principal* del sistema $\mathbf{x}' = A\mathbf{x}$.

Queda pendiente un problema mayor (y obvio). ¿Cómo calculamos e^{At} de una manera práctica? Comencemos con dos ejemplos.

Ejemplo 1 Sea $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$. Entonces

$$A^2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2^2 & 0 \\ 0 & 0 & 3^2 \end{pmatrix}, \quad A^3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2^3 & 0 \\ 0 & 0 & 3^3 \end{pmatrix}, \ldots, \quad A^m = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2^m & 0 \\ 0 & 0 & 3^m \end{pmatrix}$$

y

$$e^{At} = I + At + \frac{A^2t^2}{2!} + \frac{A^3t^3}{3!} + \cdots = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} t & 0 & 0 \\ 0 & 2t & 0 \\ 0 & 0 & 3t \end{pmatrix}$$

$$+ \begin{pmatrix} \frac{t^2}{2!} & 0 & 0 \\ 0 & \frac{2^2t^2}{2!} & 0 \\ 0 & 0 & \frac{3^2t^2}{2!} \end{pmatrix} + \begin{pmatrix} \frac{t^3}{3!} & 0 & 0 \\ 0 & \frac{2^3t^3}{3!} & 0 \\ 0 & 0 & \frac{3^3t^3}{3!} \end{pmatrix} + \cdots$$

$$= \begin{pmatrix} 1 + t + \frac{t^2}{2!} + \frac{t^3}{3!} + \cdots & 0 & 0 \\ 0 & 1 + (2t) + \frac{(2t)^2}{2!} + \frac{(2t)^3}{3!} + \cdots & 0 \\ 0 & 0 & 1 + (3t) + \frac{(3t)^2}{2!} + \frac{(3t)^3}{3!} + \cdots \end{pmatrix}$$

$$= \begin{pmatrix} e^t & 0 & 0 \\ 0 & e^{2t} & 0 \\ 0 & 0 & e^{3t} \end{pmatrix}$$

Ejemplo 2 Sea $A = \begin{pmatrix} a & 1 \\ 0 & a \end{pmatrix}$. Entonces, como es fácil de verificar,

$$A^2 = \begin{pmatrix} a^2 & 2a \\ 0 & a^2 \end{pmatrix}, A^3 = \begin{pmatrix} a^3 & 3a^2 \\ 0 & a^3 \end{pmatrix}, \ldots, A^m = \begin{pmatrix} a^m & ma^{m-1} \\ 0 & a^m \end{pmatrix}, \ldots$$

por lo que

$$e^{At} = \begin{pmatrix} \sum_{m=0}^{\infty} \frac{(at)^m}{m!} & \sum_{m=1}^{\infty} \frac{ma^{m-1}t^m}{m!} \\ 0 & \sum_{m=0}^{\infty} \frac{(at)^m}{m!} \end{pmatrix}$$

Ahora

$$\sum_{m=1}^{\infty} \frac{ma^{m-1}t^m}{m!} = \sum_{m=1}^{\infty} \frac{a^{m-1}t^m}{(m-1)!} = t + at^2 + \frac{a^2t^3}{2!} + \frac{a^3t^4}{3!} + \cdots$$

$$= t\left(1 + at + \frac{a^2t^2}{2!} + \frac{a^3t^3}{3!} + \cdots\right) = te^{at}$$

En consecuencia,

$$e^{At} = \begin{pmatrix} e^{at} & te^{at} \\ 0 & e^{at} \end{pmatrix}$$

Como lo muestra el Ejemplo 1, es fácil calcular e^{At} si A es una matriz diagonal. El Ejemplo 1 muestra que si $D = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$, entonces $e^{Dt} = \text{diag}(e^{\lambda_1 t}, e^{\lambda_2 t}, \ldots, e^{\lambda_n t})$.

En el Ejemplo 2, calculamos e^{At} para una matriz A en la forma canónica de Jordan. Resulta claro, pues, que esto es todo lo que necesitamos hacer, tal como lo sugiere el siguiente teorema.

Teorema 2 Sea J la forma canónica de Jordan de una matriz A y sea $J = C^{-1}AC$. Entonces $A = CJC^{-1}$ y

$$e^{At} = Ce^{Jt}C^{-1} \tag{16}$$

Demostración Primero notemos que

$$A^n = (CJC^{-1})^n = \overbrace{(CJC^{-1})(CJC^{-1}) \cdots (CJC^{-1})}^{n \text{ veces}}$$
$$= CJ(C^{-1}C)J(C^{-1}C)J(C^{-1}C) \cdots (C^{-1}C)JC^{-1}$$
$$= CJ^nC^{-1}$$

De aquí se deduce que

$$(At)^n = C(Jt)^nC^{-1} \tag{17}$$

Por lo tanto,

$$e^{At} = I + (At) + \frac{(At)^2}{2!} + \cdots = CIC^{-1} + C(Jt)C^{-1} + C\frac{(Jt)^2}{2!}C^{-1} + \cdots$$

$$= C\left[I + (Jt) + \frac{(Jt)^2}{2!} + \cdots\right]C^{-1} = Ce^{Jt}C^{-1} \quad ■$$

El Teorema 2 nos dice que para evaluar e^{At} lo único que realmente necesitamos calcular es e^{Jt}. Cuando J es una diagonal (como sucede en la mayoría de los casos, sabemos entonces cómo determinar e^{Jt}. Si A es una matriz de 2×2 que no es diagonalizable, entonces $J = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$ y $e^{Jt} = \begin{pmatrix} e^{\lambda t} & te^{\lambda t} \\ 0 & e^{\lambda t} \end{pmatrix}$ como lo calculamos en el Ejemplo 2. De hecho no es difícil evaluar e^{Jt} donde J es cualquier matriz de Jordan. Primero es necesario calcular e^{Bt} para una matriz bloque de Jordan B. Un método para lograrlo se da en los Problemas del 20 al 22.

Apliquemos ahora nuestros cálculos a un simple modelo biológico de crecimiento poblacional. Suponga que en un ecosistema existen dos especies S_1 y S_2 relacionadas entre sí. Denotamos las poblaciones de las especies en el momento t por $x_1(t)$ y $x_2(t)$. Un sistema que gobierna el crecimiento relativo de las dos especies es

$$\begin{aligned} x_1'(t) &= ax_1(t) + bx_2(t) \\ x_2'(t) &= cx_1(t) + dx_2(t) \end{aligned} \tag{18}$$

Podemos interpretar las constantes a, b, c y d del siguiente modo. Si las especies están en competencia, entonces es razonable tener $b<0$ y $c<0$. Esto es verdadero porque al incrementarse la población de una de las especies disminuirá el crecimiento de la otra. Un segundo modelo es una relación entre *depredador y presa*. Si S_1 es la presa y S_2 es el depredador (S_2 se come a S_1), entonces es razonable tener $b<0$ y $c>0$ puesto que un incremento en la especie *depredadora* causará una disminución en la especie *de la presa*, mientras que un incremento en la especie *de la presa* causará un incremento en la especie *depredadora* (puesto que tendrán más alimento). Finalmente, en una relación *simbiótica* (cada especie vive de la otra) tendríamos $b>0$ y $c>0$. Por supuesto, las constantes a, b, c y d dependen de una gran variedad de factores incluyendo alimento disponible, temporada del año, clima, límites debido a la sobrepoblación, competencia con otras especies y muchas otras. Analizaremos cuatro modelos diferentes usando el material en esta sección. Supongamos que t está medido en años.

Ejemplo 3 ***Un modelo de especies en competencia.*** Considere el sistema

$$\begin{aligned} x_1'(t) &= 3x_1(t) - x_2(t) \\ x_2'(t) &= -2x_1(t) + 2x_2(t) \end{aligned}$$

Este sistema describe la influencia de las poblaciones de dos especies en competencia, sobre sus respectivas tasas de crecimiento. Suponga que las poblaciones iniciales son $x_1(0) = 90$ y $x_2(0) = 150$. Determine las poblaciones de ambas especies para $t > 0$.

Solución Tenemos $A=\begin{pmatrix}3 & -1\\ -2 & 2\end{pmatrix}$. Los valores característicos de A son $\lambda_1=1$ y $\lambda_2=4$ con sus correspondientes vectores característicos $\mathbf{v}_1=\begin{pmatrix}1\\2\end{pmatrix}$ y $\mathbf{v}_2=\begin{pmatrix}1\\-1\end{pmatrix}$. Entonces

$$C=\begin{pmatrix}1 & 1\\ 2 & -1\end{pmatrix}\quad C^{-1}=-\frac{1}{3}\begin{pmatrix}-1 & -1\\ -2 & 1\end{pmatrix}\quad J=D=\begin{pmatrix}1 & 0\\ 0 & 4\end{pmatrix}\quad e^{Jt}=\begin{pmatrix}e^t & 0\\ 0 & e^{4t}\end{pmatrix}$$

$$e^{At}=Ce^{Jt}C^{-1}=-\frac{1}{3}\begin{pmatrix}1 & 1\\ 2 & -1\end{pmatrix}\begin{pmatrix}e^t & 0\\ 0 & e^{4t}\end{pmatrix}\begin{pmatrix}-1 & -1\\ -2 & 1\end{pmatrix}$$

$$=-\frac{1}{3}\begin{pmatrix}1 & 1\\ 2 & -1\end{pmatrix}\begin{pmatrix}-e^t & -e^t\\ -2e^{4t} & e^{4t}\end{pmatrix}$$

$$=-\frac{1}{3}\begin{pmatrix}-e^t-2e^{4t} & -e^t+e^{4t}\\ -2e^t+2e^{4t} & -2e^t-e^{4t}\end{pmatrix}$$

Finalmente, la solución al sistema está dado por

$$\mathbf{x}(t)=\begin{pmatrix}x_1(t)\\ x_2(t)\end{pmatrix}=e^{At}\mathbf{x}_0=-\frac{1}{3}\begin{pmatrix}-e^t-2e^{4t} & -e^t+e^{4t}\\ -2e^t+2e^{4t} & -2e^t-e^{4t}\end{pmatrix}\begin{pmatrix}90\\ 150\end{pmatrix}$$

$$=-\frac{1}{3}\begin{pmatrix}-240e^t-30e^{4t}\\ -480e^t+30e^{4t}\end{pmatrix}=\begin{pmatrix}80e^t+10e^{4t}\\ 160e^t-10e^{4t}\end{pmatrix}$$

Por ejemplo, después de 6 meses ($t=\frac{1}{2}$ año), $x_1(t)=80e^{1/2}+10e^2\approx 206$ individuos, mientras que $x_2(t)=160e^{1/2}-10e^2\approx 190$ individuos. Más significativo aún es $160e^t - 10e^{4t} = 0$ cuando $16e^t = e^{4t}$ o $16 = e^{3t}$ o $3t = \ln 16$ y $t=(\ln 16)/3\approx 2.77/3\approx 0.92$ años ≈ 11 meses. De esta manera la segunda especie será eliminada después de escasamente 11 meses, aun cuando comenzó con una población más grande. En los Problemas 10 y 11 se le pide mostrar que ninguna población será eliminada si $x_2(0)=2x_1(0)$ y que la primera población será eliminada si $x_2(0) > 2x_1(0)$. Por lo tanto, como Darwin sabía muy bien, la supervivencia en este modelo muy simple depende de los tamaños relativos de las especies competitivas cuando comienza la competencia.

Ejemplo 4 ***Un modelo depredador-presa.*** Consideremos el siguiente sistema en el que la especie 1 es la presa y la especie 2 es el depredador:

$$x_1'(t)=2x_1(t)-x_2(t)$$
$$x_2'(t)=x_1(t)+4x_2(t)$$

Encuentre las poblaciones de dos especies para $t>0$ si las poblaciones iniciales son $x_1(0)=500$ y $x_2(0)=100$.

Solución Aquí $A=\begin{pmatrix}2 & -1\\ 1 & 4\end{pmatrix}$ y el único valor característico es $\lambda=3$ con el único vector característico $\begin{pmatrix}1\\-1\end{pmatrix}$. Una solución a la ecuación $(A-3I)\mathbf{v}_2=\mathbf{v}_1$ (vea Teorema 6.6.2) es $\mathbf{v}_2=\begin{pmatrix}1\\-2\end{pmatrix}$. Entonces

$$C=\begin{pmatrix}1 & 1\\ -1 & -2\end{pmatrix}\quad C^{-1}=\begin{pmatrix}2 & 1\\ -1 & -1\end{pmatrix}\quad J=\begin{pmatrix}3 & 1\\ 0 & 3\end{pmatrix}$$

$$e^{Jt}=\begin{pmatrix}e^{3t} & te^{3t}\\ 0 & e^{3t}\end{pmatrix}=e^{3t}\begin{pmatrix}1 & t\\ 0 & 1\end{pmatrix}\quad \text{(del Ejemplo 2)}$$

y

$$e^{At}=Ce^{Jt}C^{-1}=e^{3t}\begin{pmatrix}1 & 1\\ -1 & -2\end{pmatrix}\begin{pmatrix}1 & t\\ 0 & 1\end{pmatrix}\begin{pmatrix}2 & 1\\ -1 & -1\end{pmatrix}$$

$$=e^{3t}\begin{pmatrix}1 & 1\\ -1 & -2\end{pmatrix}\begin{pmatrix}2-t & 1-t\\ -1 & -1\end{pmatrix}$$

$$=e^{3t}\begin{pmatrix}1-t & -t\\ t & 1+t\end{pmatrix}$$

De esta manera la solución al sistema es

$$\mathbf{x}(t)=\begin{pmatrix}x_1(t)\\ x_2(t)\end{pmatrix}=e^{At}\mathbf{x}_0=e^{3t}\begin{pmatrix}1-t & -t\\ t & 1+t\end{pmatrix}\begin{pmatrix}500\\ 100\end{pmatrix}=e^{3t}\begin{pmatrix}500-600t\\ 100+600t\end{pmatrix}$$

Es obvio que la especie presa será eliminada despues de $\frac{5}{6}$ año = 10 meses, aun cuando comenzó con una población cinco veces más grande que la especie depredadora. De hecho es fácil mostrar (Problema 12) que independientemente del tamaño inicial de la especie presa, ésta será eliminada en menos de 1 año.

Ejemplo 5 ***Otro modelo depredador-presa.*** Consideremos el modelo depredador-presa regido por el sistema.

$$x_1'(t)=x_1(t)+x_2(t)$$
$$x_2'(t)=-x_1(t)+x_2(t)$$

Si las poblaciones iniciales son $x_1(0) = x_2(0) = 1000$, determine las poblaciones de las dos especies para $t > 0$.

Solución Aquí $A=\begin{pmatrix}1 & 1\\ -1 & 1\end{pmatrix}$ con ecuación característica $\lambda^2-2\lambda+2=0$, raíces complejas $\lambda_1 = 1+i$ y $\lambda_2 = 1-i$ y vectores característicos $\mathbf{v}_1=\begin{pmatrix}1\\ i\end{pmatrix}$ y $\mathbf{v}_2=\begin{pmatrix}1\\ -i\end{pmatrix}$.*

Entonces

$$C=\begin{pmatrix}1 & 1\\ i & -i\end{pmatrix}\quad C^{-1}=-\frac{1}{2i}\begin{pmatrix}-i & -1\\ -i & 1\end{pmatrix}=\frac{1}{2}\begin{pmatrix}1 & -i\\ 1 & i\end{pmatrix}\quad J=D=\begin{pmatrix}1+i & 0\\ 0 & 1-i\end{pmatrix}$$

y

$$e^{Jt}=\begin{pmatrix}e^{(1+i)t} & 0\\ 0 & e^{(1-i)t}\end{pmatrix}$$

---

* Nótese que $\lambda_2=\overline{\lambda_1}$ y $\mathbf{v}_2=\overline{\mathbf{v}_1}$. Esto no debe sorprender, porque de acuerdo con el resultado del Problema 6.1.33, los valores propios de matrices reales aparecen en pares complejos conjugados y sus vectores propios correspondientes son complejos conjugados.

Ahora, por la fórmula de Euler (Apéndice 2), $e^{it} = \cos t + i \operatorname{sen} t$. De esta manera $e^{(1+i)t} = e^t e^{it} = e^t(\cos t + i \operatorname{sen} t)$. Análogamente, $e^{(1-i)t} = e^t e^{-it} = e^t(\cos t - i \operatorname{sen} t)$.

En consecuencia

$$e^{Jt} = e^t\begin{pmatrix}\cos t + i \operatorname{sen} t & 0 \\ 0 & \cos t - i \operatorname{sen} t\end{pmatrix}$$

y

$$e^{At} = Ce^{Jt}C^{-1} = \frac{e^t}{2}\begin{pmatrix}1 & 1 \\ i & -i\end{pmatrix}\begin{pmatrix}\cos t + i \operatorname{sen} t & 0 \\ 0 & \cos t - i \operatorname{sen} t\end{pmatrix}\begin{pmatrix}1 & -i \\ 1 & i\end{pmatrix}$$

$$= \frac{e^t}{2}\begin{pmatrix}1 & 1 \\ i & -i\end{pmatrix}\begin{pmatrix}\cos t + i \operatorname{sen} t & -i\cos t + \operatorname{sen} t \\ \cos t - i \operatorname{sen} t & i\cos t + \operatorname{sen} t\end{pmatrix}$$

$$= \frac{e^t}{2}\begin{pmatrix}2\cos t & 2 \operatorname{sen} t \\ -2 \operatorname{sen} t & 2\cos t\end{pmatrix} = e^t\begin{pmatrix}\cos t & \operatorname{sen} t \\ -\operatorname{sen} t & \cos t\end{pmatrix}$$

Finalmente,

$$\mathbf{x}(t) = e^{At}\mathbf{x}(0) = e^t\begin{pmatrix}\cos t & \operatorname{sen} t \\ -\operatorname{sen} t & \cos t\end{pmatrix}\begin{pmatrix}1000 \\ 1000\end{pmatrix} = \begin{pmatrix}1000e^t(\cos t + \operatorname{sen} t) \\ 1000e^t(\cos t - \operatorname{sen} t)\end{pmatrix}$$

La especie presa comienza a extinguirse cuando $1000e^t(\cos t - \operatorname{sen} t) = 0$ o cuando $\operatorname{sen} t = \cos t$. Se tiene que la primera solución positiva de esta última ecuación es $t = \pi/4 \approx 0.7854$ año ≈ 9.4 meses.

Ejemplo 6 ***Un modelo de cooperación entre especies (Simbiosis).*** Consideremos el modelo simbiótico gobernado por el sistema

$$x_1'(t) = -\tfrac{1}{2}x_1(t) + x_2(t)$$
$$x_2'(t) = \tfrac{1}{4}x_1(t) - \tfrac{1}{2}x_2(t)$$

Note que en este modelo la población de cada especie incrementa proporcionalmente a la población de la otra y decrece proporcionalmente con respecto a su propia población. Suponga que $x_1(0) = 200$ y $x_2(0) = 500$. Determine la población de cada especie para $t > 0$.

Solución Aquí $A = \begin{pmatrix}-\frac{1}{2} & 1 \\ \frac{1}{4} & -\frac{1}{2}\end{pmatrix}$ con valores característicos $\lambda_1 = 0$ y $\lambda_2 = -1$ y sus correspondientes vectores característicos $\mathbf{v}_1 = \begin{pmatrix}2 \\ 1\end{pmatrix}$ y $\mathbf{v}_2 = \begin{pmatrix}2 \\ -1\end{pmatrix}$. Entonces

$$C = \begin{pmatrix}2 & 2 \\ 1 & -1\end{pmatrix}, \quad C^{-1} = -\tfrac{1}{4}\begin{pmatrix}-1 & -2 \\ -1 & 2\end{pmatrix}, \quad J = D = \begin{pmatrix}0 & 0 \\ 0 & -1\end{pmatrix},$$

y $e^{Jt} = \begin{pmatrix}e^{0t} & 0 \\ 0 & e^{-t}\end{pmatrix} = \begin{pmatrix}1 & 0 \\ 0 & e^{-t}\end{pmatrix}$. De esta manera

$$e^{At} = -\frac{1}{4}\begin{pmatrix}2 & 2 \\ 1 & -1\end{pmatrix}\begin{pmatrix}1 & 0 \\ 0 & e^{-t}\end{pmatrix}\begin{pmatrix}-1 & -2 \\ -1 & 2\end{pmatrix}$$

$$= -\frac{1}{4}\begin{pmatrix}2 & 2 \\ 1 & -1\end{pmatrix}\begin{pmatrix}-1 & -2 \\ -e^{-t} & 2e^{-t}\end{pmatrix}$$

$$= -\frac{1}{4}\begin{pmatrix}-2 - 2e^{-t} & -4 + 4e^{-t} \\ -1 + e^{-t} & -2 - 2e^{-t}\end{pmatrix}$$

y así

$$\mathbf{x}(t) = e^{At}\mathbf{x}(0) = -\frac{1}{4}\begin{pmatrix}-2 - 2e^{-t} & -4 + 4e^{-t} \\ -1 + e^{-t} & -2 - 2e^{-t}\end{pmatrix}\begin{pmatrix}200 \\ 500\end{pmatrix}$$

$$= -\frac{1}{4}\begin{pmatrix}-2400 + 1600e^{-t} \\ -1200 - 800e^{-t}\end{pmatrix}$$

$$= \begin{pmatrix}600 - 400e^{-t} \\ 300 + 200e^{-t}\end{pmatrix}$$

Note que $e^{-t} \to 0$ cuando $t \to \infty$. Esto significa que conforme pase el tiempo, las dos especies simbióticas se aproximan a las poblaciones de *equilibrio* de 600 y 300, respectivamente. Ninguna población es eliminada.

Problemas 6.7

En los Problemas del 1 al 9 encuentre la matriz solución principal e^{At} del sistema $\mathbf{x}'(t) = A\mathbf{x}(t)$.

1. $A = \begin{pmatrix}-2 & -2 \\ -5 & 1\end{pmatrix}$
2. $A = \begin{pmatrix}3 & -1 \\ -2 & 4\end{pmatrix}$
3. $A = \begin{pmatrix}2 & -1 \\ 5 & -2\end{pmatrix}$
4. $A = \begin{pmatrix}3 & -5 \\ 1 & -1\end{pmatrix}$
5. $A = \begin{pmatrix}-10 & -7 \\ 7 & 4\end{pmatrix}$
6. $A = \begin{pmatrix}-2 & 1 \\ 5 & 2\end{pmatrix}$
7. $A = \begin{pmatrix}-12 & 7 \\ -7 & 2\end{pmatrix}$
8. $A = \begin{pmatrix}1 & 1 & -2 \\ -1 & 2 & 1 \\ 0 & 1 & -1\end{pmatrix}$
9. $A = \begin{pmatrix}4 & 6 & 6 \\ 1 & 3 & 2 \\ -1 & -5 & -2\end{pmatrix}$

10. En el Ejemplo 3 muestre que si el vector inicial $\mathbf{x}(0) = \begin{pmatrix}a \\ 2a\end{pmatrix}$, donde a es una constante, entonces ambas poblaciones crecen a una tasa proporcional a e^t.
11. En el Ejemplo 3, muestre que si $x_2(0) > 2x_1(0)$, entonces la primera población será eliminada.
12. En el Ejemplo 4, muestre que la primera población se extinguirá en α años, donde $a = x_1(0)/[x_1(0) + x_2(0)]$.

★ 13. En una planta desalinizadora de agua hay dos tanques de agua. Suponga que el tanque 1 contiene 1000 litros de salmuera en el cual están disueltos 1000 kg de sal, y el tanque 2 contiene 1000 litros de agua pura. Suponga que el agua fluye hacia el tanque 1 a una razón de 20 litros por minuto y la mezcla fluye del tanque 1 al tanque 2 a una razón de 30 litros por minuto. Se bombean 10 litros del tanque 2 al tanque 1 (estableciendo *retroalimentación*), mientras que se extraen 20 litros del tanque. Encuentre la cantidad de sal en ambos tanques en cualquier instante t. [*Su-*

gerencia: Escriba la información como un sistema de 2 × 2 y denotemos con $x_1(t)$ y $x_2(t)$ las cantidades de sal en cada tanque.]

14. Una comunidad de n individuos es expuesta a una enfermedad infecciosa.* En un momento t dado, la comunidad es dividida en tres grupos: el grupo 1 con una población $x_1(t)$ es el grupo susceptible; el grupo 2 con una población $x_2(t)$ es el grupo de individuos infectados en circulación; y el grupo 3 de población $x_3(t)$, está formado por los que han sido aislados, que se han muerto o que han sido inmunizados. Es razonable suponer que inicialmente $x_2(t)$ y $x_3(t)$ serán pequeñas comparadas con $x_1(t)$. Sean α y β constantes positivas que denotan la proporción en que se infectan los susceptibles y los infectados pasan a formar parte del grupo 3, respectivamente. Entonces un modelo razonable para la propagación de la enfermedad está dada por el sistema

$$\begin{aligned} x_1'(t) &= -\alpha x_1(0)x_2 \\ x_2'(t) &= \alpha x_1(0)x_2 - \beta x_2 \\ x_3'(t) &= \beta x_2 \end{aligned}$$

a. Escriba este sistema en la forma $\mathbf{x}' = A\mathbf{x}$ y encuentre la solución en términos de $x_1(0)$, $x_2(0)$ y $x_3(0)$. Note que $x_1(0) + x_2(0) + x_3(0) = n$.

b. Muestre que si $\alpha x(0) < \beta$, entonces la enfermedad no producirá una epidemia.

c. ¿Qué sucederá si $\alpha x(0) > \beta$?

15. Considere la *ecuación diferencial de segundo orden* $x''(t) + ax'(t) + bx(t) = 0$.

a. Haciendo $x_1(t) = x(t)$ y $x_2(t) = x'(t)$, escriba la ecuación anterior como un sistema de primer orden en la forma de la Ecuación (7), donde A es una matriz de 2 × 2.

b. Muestre que la ecuación característica de A es $\lambda^2 + a\lambda + b = 0$.

En los Problemas del 16 *al* 19 *use el resultado del Problema* 15 *para resolver la ecuación dada*

16. $x'' + 5x' + 6x = 0$; $x(0) = 1$, $x'(0) = 0$

17. $x'' + 6x' + 9x = 0$; $x(0) = 1$, $x'(0) = 2$

18. $x'' + 4x = 0$; $x(0) = 0$, $x'(0) = 1$

19. $x'' - 3x' - 10x = 0$; $x(0) = 3$, $x'(0) = 2$

20. Sea $N_3 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$. Muestre que $N_3^3 = 0$, es la matriz cero.

21. Muestre que $e^{N_3t} = \begin{pmatrix} 1 & t & t^2/2 \\ 0 & 1 & t \\ 0 & 0 & 1 \end{pmatrix}$. [*Sugerencia*: Escriba la serie para e^{N_3t} y use el resultado del Problema 20.]

22. Sea $J = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}$. Muestre que $e^{Jt} = e^{\lambda t}\begin{pmatrix} 1 & t & t^2/2 \\ 0 & 1 & t \\ 0 & 0 & 1 \end{pmatrix}$. [*Sugerencia:* $Jt = \lambda It + N_3t$. Use el hecho de que $e^{A+B} = e^Ae^B$.]

23. Usando el resultado del Problema 22, calcule e^{At}, donde $A = \begin{pmatrix} -2 & 1 & 0 \\ -2 & 1 & -1 \\ -1 & 1 & -2 \end{pmatrix}$. [*Sugerencia:* Vea el Problema 6.6.20.]

* Una discusión de este modelo aparece en "The Total Size of a General Stochastic Epidemic", de N. Bailey, *Biometrika* 40(1953): 177-185.

24. Calcule e^{At}, donde $A = \begin{pmatrix} -1 & -18 & -7 \\ 1 & -13 & -4 \\ -1 & 25 & 8 \end{pmatrix}$.

25. Calcule e^{Jt}, donde $J = \begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{pmatrix}$.

26. Calcule e^{At}, donde $A = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 3 \end{pmatrix}$.

27. Calcule e^{At}, donde $A = \begin{pmatrix} -4 & 1 & 0 & 0 \\ 0 & -4 & 1 & 0 \\ 0 & 0 & -4 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}$.

6.8 Una perspectiva diferente: Los teoremas de Cayley-Hamilton y Gershgorin

Existen muchos resultados interesantes relativos a los valores característicos de una matriz. En esta sección vamos a discutir dos de los más útiles. El primero dice que cualquier matriz satisface su propia ecuación característica. El segundo muestra cómo localizar los valores característicos de cualquier matriz prácticamente sin hacer cálculos.

Sea $p(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$ un polinomio y sea A una matriz de $n \times n$. Entonces las potencias de A están definidas y con ellas definimos

$$p(A) = A^n + a_{n-1}A^{n-1} + \cdots + a_1A + a_0I \tag{1}$$

Ejemplo 1 Sea $A = \begin{pmatrix} -1 & 4 \\ 3 & 7 \end{pmatrix}$ y $p(x) = x^2 - 5x + 3$. Entonces

$$P(A) = A^2 - 5A + 3I = \begin{pmatrix} 13 & 24 \\ 18 & 61 \end{pmatrix} + \begin{pmatrix} 5 & -20 \\ -15 & -35 \end{pmatrix} + \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix} = \begin{pmatrix} 21 & 4 \\ 3 & 29 \end{pmatrix}$$

La expresión (1) es un polinomio con coeficientes escalares definidos por una matriz variable. También podemos definir un polinomio con coeficientes *matriciales cuadrados* como

$$Q(\lambda) = B_0 + B_1\lambda + B_2\lambda^2 + \cdots + B_n\lambda^n \tag{2}$$

Si A es una matriz, entonces definimos

$$Q(A) = B_0 + B_1A + B_2A^2 + \cdots + B_mA^m \tag{3}$$

Debemos tener cuidado en (3), puesto que las matrices no son conmutativas para la multiplicación.

Teorema 1 Si $P(\lambda)$ y $Q(\lambda)$ son polinomios en la variable escalar λ, con coeficientes matriciales cuadrados y si $P(\lambda)=Q(\lambda)(A-\lambda I)$, entonces $P(A) = 0$.

Demostración Si $Q(\lambda)$ está dado por la Ecuación (2), entonces

$$\begin{aligned}P(\lambda) &= (B_0+B_1\lambda+B_2\lambda^2+\cdots+B_n\lambda^n)(A-\lambda I)\\ &= B_0A+B_1A\lambda+B_2A\lambda^2+\cdots+B_nA\lambda^n\\ &\quad -B_0\lambda-B_1\lambda^2-B_2\lambda^3-\cdots-B_n\lambda^{n+1}\end{aligned} \tag{4}$$

Entonces, sustituyendo λ por A en (4), obtenemos

$$\begin{aligned}P(A) &= B_0A+B_1A^2+B_2A^3+\cdots+B_nA^{n+1}\\ &\quad -B_0A-B_1A^2-B_2A^3-\cdots-B_nA^{n+1}=0 \quad ■\end{aligned}$$

Nota. No podemos demostrar este teorema sustituyendo $\lambda = A$ para obtener $P(A) = Q(A)(A-A) = 0$ porque es posible encontrar polinomios $P(\lambda)$ y $Q(\lambda)$ con coeficientes matriciales tales que $F(\lambda) = P(\lambda)Q(\lambda)$ pero $F(A) \neq P(A)Q(A)$. (Vea el Problema 17.)

Con esto ya podemos demostrar nuestro primer teorema importante.

Teorema 2 ***El Teorema de Cayley-Hamilton.**** Cada matriz cuadrada satisface su propia ecuación característica. Esto es, si $p(\lambda) = 0$ es la ecuación característica de A, entonces $p(A) = 0$

Demostración Tenemos

$$p(\lambda)=\det(A-\lambda I)=\begin{vmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22}-\lambda & \cdots & a_{2n}\\ \vdots & \vdots & & \vdots\\ a_{n1} & a_{n2} & \cdots & a_{nn}-\lambda\end{vmatrix}$$

Claramente, cualquier cofactor de $(A-\lambda I)$ es un polinomio en λ. Por lo tanto la adjunta de $A-\lambda I$ (vea Definición 2.4.1) es una matriz de $n \times n$, en la que cada una de sus componentes es un polinomio en λ. Esto es:

$$\operatorname{adj}(A-\lambda I)=\begin{pmatrix} p_{11}(\lambda) & p_{12}(\lambda) & \cdots & p_{1n}(\lambda)\\ p_{21}(\lambda) & p_{22}(\lambda) & \cdots & p_{2n}(\lambda)\\ \vdots & \vdots & & \vdots\\ p_{n1}(\lambda) & p_{22}(\lambda) & \cdots & p_{nn}(\lambda)\end{pmatrix}$$

* Llamado así en honor a Sir William Rowan Hamilton (de quien se habló en el Capítulo 1) y a Arthur Cayley (1821-1895). Cayley publicó la primera discusión de este famoso teorema en 1858. Hamilton descubrió el resultado independientemente en su trabajo sobre los cuaterniones.

Esto significa que podemos considerar adj $(A-\lambda I)$ como un polinomio $Q(\lambda)$, en λ, con coeficientes matriciales de $n\times n$. Para ver esto, observe lo siguiente:

$$\begin{pmatrix}-\lambda^2-2\lambda+1 & 2\lambda^2-7\lambda-4\\ 4\lambda^2+5\lambda-2 & -3\lambda^2-\lambda+3\end{pmatrix}=\begin{pmatrix}-1 & 2\\ 4 & -3\end{pmatrix}\lambda^2+\begin{pmatrix}-2 & -7\\ 5 & -1\end{pmatrix}\lambda+\begin{pmatrix}1 & -4\\ -2 & 3\end{pmatrix}$$

Ahora, del Teorema 2.4.2,

$$\det(A-\lambda I)I=[\operatorname{adj}(A-\lambda I)][A-\lambda I]=Q(\lambda)(A-\lambda I) \tag{5}$$

Pero $\det(A-\lambda I)I=p(\lambda)I$. Si

$$p(\lambda)=\lambda^n+a_{n-1}\lambda^{n-1}+\cdots+a_1\lambda+a_0,$$

entonces definimos

$$P(\lambda)=p(\lambda)I=\lambda^nI+a_{n-1}\lambda^{n-1}I+\cdots+a_1\lambda I+a_0I.$$

De esta manera, de (5), tenemos $P(\lambda)=Q(\lambda)(A-\lambda I)$. Finalmente, del Teorema 1, $P(A)=0$. Esto completa la demostración. ■

Ejemplo 2 Sea $A = \begin{pmatrix}1 & -1 & 4\\ 3 & 2 & -1\\ 2 & 1 & -1\end{pmatrix}$. En el Ejemplo 6.1.4 calculamos la ecuación característica $\lambda^3 - 2\lambda^2 - 5\lambda + 6 = 0$. Ahora calculamos

$$A^2=\begin{pmatrix}6 & 1 & 1\\ 7 & 0 & 11\\ 3 & -1 & 8\end{pmatrix},\quad A^3=\begin{pmatrix}11 & -3 & 22\\ 29 & 4 & 17\\ 16 & 3 & 5\end{pmatrix}$$

y

$$\begin{aligned}A^3-2A^2-5A+6I &= \begin{pmatrix}11 & -3 & 22\\ 29 & 4 & 17\\ 16 & 3 & 5\end{pmatrix}+\begin{pmatrix}-12 & -2 & -2\\ -14 & 0 & -22\\ -6 & 2 & -16\end{pmatrix}\\ &\quad+\begin{pmatrix}-5 & 5 & -20\\ -15 & -10 & 5\\ -10 & -5 & 5\end{pmatrix}+\begin{pmatrix}6 & 0 & 0\\ 0 & 6 & 0\\ 0 & 0 & 6\end{pmatrix}\\ &= \begin{pmatrix}0 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}\end{aligned}$$

En algunos casos el teorema de Cayley-Hamilton sirve para calcular la inversa de una matriz. Si A^{-1} existe y $p(A) = 0$, se tiene que $A^{-1}p(A) = 0$. $p(\lambda)=\lambda^n+a_{n-1}\lambda^{n-1}+\cdots+a_1\lambda+a_0$, entonces

$$p(A)=A^n+a_{n-1}A^{n-1}+\cdots+a_1A+a_0I=0$$

y

$$A^{-1}p(A)=A^{n-1}+a_{n-1}A^{n-2}+\cdots+a_2A+a_1I+a_0A^{-1}=0$$

En consecuencia

$$A^{-1}=\frac{1}{a_0}(-A^{n-1}-a_{n-1}A^{n-2}-\cdots-a_2A-a_1I) \qquad (6)$$

Note que $a_0\neq 0$ porque $a_0=\det A$ (¿Por qué?) y supusimos que A era invertible.

Ejemplo 3 Sea $A=\begin{pmatrix}1 & -1 & 4\\ 3 & 2 & -1\\ 2 & 1 & -1\end{pmatrix}$. Entonces $p(\lambda)=\lambda^3-2\lambda^2-5\lambda+6$. Aquí $n=3$, $a_2=-2$, $a_1=-5$, $a_0=6$, y

$$A^{-1}=\frac{1}{6}(-A^2+2A+5I)$$

$$=\frac{1}{6}\left[\begin{pmatrix}-6 & -1 & -1\\ -7 & 0 & -11\\ -3 & 1 & -8\end{pmatrix}+\begin{pmatrix}2 & -2 & 8\\ 6 & 4 & -2\\ 4 & 2 & -2\end{pmatrix}+\begin{pmatrix}5 & 0 & 0\\ 0 & 5 & 0\\ 0 & 0 & 5\end{pmatrix}\right]$$

$$=\frac{1}{6}\begin{pmatrix}1 & -3 & 7\\ -1 & 9 & -13\\ 1 & 3 & -5\end{pmatrix}$$

Note que calculamos A^{-1} con una sola división y calculando únicamente un determinante (para encontrar $p(\lambda)=\det(A-\lambda I)$). Este método es a veces muy eficiente en una computadora.

Ahora veamos el segundo resultado importante de esta sección. Sea A una matriz de $n\times n$. Escribimos, como siempre

$$A=\begin{pmatrix}a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & & \vdots\\ a_{n1} & a_{n2} & \cdots & a_{nn}\end{pmatrix}$$

Definimos el número

$$r_1=|a_{12}|+|a_{13}|+\cdots+|a_{1n}|=\sum_{j=2}^{n}|a_{1j}| \qquad (7)$$

Análogamente definimos

$$r_i=|a_{i1}|+|a_{i2}|+\cdots+|a_{i,i-1}|+|a_{i,i+1}|+\cdots+|a_{i,n}| = \sum_{\substack{j=1\\ j\neq i}}^{n}|a_{ij}| \qquad (8)$$

Esto es, r_i es la suma de los valores absolutos de los números en el i-ésimo renglón de A que no están en la diagonal principal de A. Sea

$$D_i=\{z\in\mathbb{C}:\ |z-a_{ii}|\le r_i\} \qquad (9)$$

Aquí D_i es un disco en el plano complejo, con centro en a_{ii}, y radio r_i (Figura 6.4). El disco D_i consiste en todos los puntos en el plano complejo dentro y sobre la circunferencia $C_i=\{z\in\mathbb{C}: |z-a_{ii}|=r_i\}$. Las circunferencias C_i, $i=1, 2, \ldots, n$, se denominan *círculos de Gershgorin*.

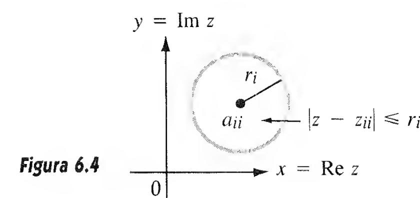


Figura 6.4

Teorema 3 ***Teorema del círculo de Gershgorin.**** Sea A una matriz de $n\times n$ y sea D_i definida por la Ecuación (9). Entonces, cada valor característico de A está contenido al menos en uno de los D_i. Esto es, si los valores característicos de A son $\lambda_1, \lambda_2, \ldots, \lambda_k$, entonces:

$$\{\lambda_1,\lambda_2,\ldots,\lambda_k\}\subset\bigcup_{i=1}^{n}D_i \qquad (10)$$

Demostración Sea λ un valor característico de A con vector característico $\mathbf{v}=\begin{pmatrix}x_1\\ x_2\\ \vdots\\ x_n\end{pmatrix}$. Sea $m=\text{máx}\{|x_1|,|x_2|,\ldots,|x_n|\}$. Entonces $(1/m)\mathbf{v}=\begin{pmatrix}y_1\\ y_2\\ \vdots\\ y_n\end{pmatrix}$ es un vector caracterís-

* El matemático ruso S. Gershgorin publicó este resultado en 1931.

tico de A correspondiente a λ y máx $\{|y_1|, |y_2|, \ldots, |y_n|\} = 1$. Sea y_i la componente de $\mathbf{y}$ con $|y_i| = 1$. Ahora $A\mathbf{y} = \lambda\mathbf{y}$. El i-ésimo componente del n-vector $A\mathbf{y}$ es $a_{i1}y_1 + a_{i2}y_2 + \cdots + a_{in}y_n$. La i-ésima componente de $\lambda\mathbf{y}$ es λy_i. De esta manera

$$a_{i1}y_1 + a_{i2}y_2 + \cdots + a_{in}y_n = \lambda y_i,$$

lo cual escribimos como

$$\sum_{j=1}^{n} a_{ij}y_j = \lambda y_i \tag{11}$$

Si restamos $a_{ii}y_i$ de ambos lados, la Ecuación (11) puede escribirse como

$$\sum_{\substack{j=1\\ j\neq i}}^{n} a_{ij}y_j = \lambda y_i - a_{ii}y_i = (\lambda - a_{ii})y_i \tag{12}$$

Después tomamos el valor absoluto de ambos lados de (12) y usando la desigualdad del triángulo $(|a+b| \le |a|+|b|)$, obtenemos

$$|(a_{ii}-\lambda)y_i| = \left| -\sum_{\substack{j=1\\ j\neq i}}^{n} a_{ij}y_j \right| \le \sum_{\substack{j=1\\ j\neq i}}^{n} |a_{ij}|\,|y_j| \tag{13}$$

Dividimos ambos lados de (13) por $|y_i|$ (el cual es igual a 1) para obtener

$$|a_{ii}-\lambda| \le \sum_{\substack{j=1\\ j\neq i}}^{n} |a_{ij}| \frac{|y_j|}{|y_i|} \le \sum_{\substack{j=1\\ j\neq i}}^{n} |a_{ij}| = r_i \tag{14}$$

El paso anterior se debió al hecho de que $|y_j| \le |y_i|$ (por la forma en que escogimos y_i). Pero esto demuestra el teorema puesto que (14) muestra que $\lambda \in D_i$. ■

Ejemplo 4 Sea $A = \begin{pmatrix} 1 & -1 & 4 \\ 3 & 2 & -1 \\ 2 & 1 & -1 \end{pmatrix}$. Entonces $a_{11}=1$, $a_{22}=2$, $a_{33}=-1$, $r_1=|-1|+|4|=5$, $r_2=|3|+|-1|=4$, y $r_3=|2|+|1|=3$. De esta manera los valores característicos de A se hallan localizados dentro de los límites de las tres circunferencias ilustradas en la Figura 6.5. Podemos verificar esto ya que sabemos por el Ejemplo 6.1.4 que los valores característicos de A son 1, -2 y 3, los cuales están localizados dentro de las tres circunferencias. Nótese que los círculos de Gershgorin pueden intersecarse.

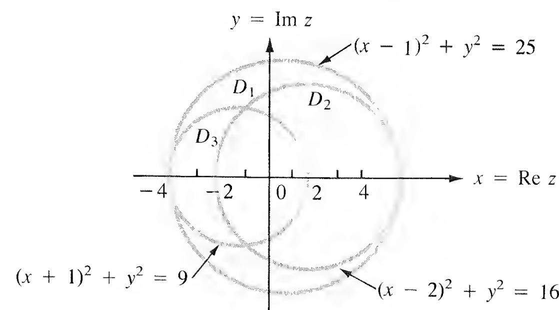


Figura 6.5

Ejemplo 5 Encuentre las cotas para los valores característicos de la matriz

$$A = \begin{pmatrix} 3 & 0 & -1 & -\frac{1}{4} & \frac{1}{4} \\ 0 & 5 & \frac{1}{2} & 0 & 1 \\ -\frac{1}{4} & 0 & 6 & \frac{1}{4} & \frac{1}{2} \\ 0 & -1 & \frac{1}{2} & -3 & \frac{1}{4} \\ \frac{1}{6} & -\frac{1}{6} & \frac{1}{3} & \frac{1}{3} & 4 \end{pmatrix}$$

Solución Aquí $a_{11}=3$, $a_{22}=5$, $a_{33}=6$, $a_{44}=-3$, $a_{55}=4$, $r_1=\frac{3}{2}$, $r_2=\frac{3}{2}$, $r_3=1$, $r_4=\frac{7}{4}$, y $r_5 = 1$. Los círculos de Gershgorin están ilustrados en la Figura 6.6. Es claro a partir del Teorema 3 y de la Figura 6.6 que si λ es un valor característico de A, entonces $|\lambda| \le 7$ y $\operatorname{Re}\lambda \ge -\frac{19}{4}$. Note la capacidad del teorema de Gershgorin para encontrar la ubicación aproximada de los valores característicos con muy poco esfuerzo.

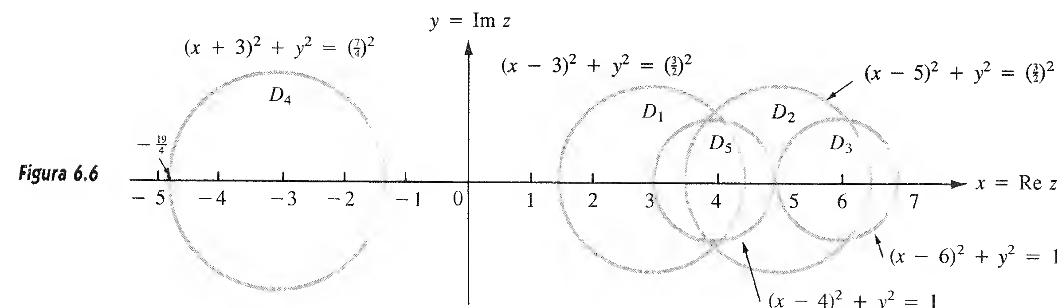


Figura 6.6

Problemas 6.8

En los Problemas del 1 al 9: **(a)** *Encuentre la ecuación característica* $p(\lambda) = 0$ *de la matriz dada;* **(b)** *verifique que* $p(A) = 0$; **(c)** *use la parte (b) para calcular* A^{-1}.

1. $\begin{pmatrix} -2 & -2 \\ -5 & 1 \end{pmatrix}$ **2.** $\begin{pmatrix} 2 & -1 \\ 5 & -2 \end{pmatrix}$ **3.** $\begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix}$

4. $\begin{pmatrix} 1 & 2 & 2 \\ 0 & 2 & 1 \\ -1 & 2 & 2 \end{pmatrix}$ **5.** $\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -3 & 3 \end{pmatrix}$ **6.** $\begin{pmatrix} -3 & -7 & -5 \\ 2 & 4 & 3 \\ 1 & 2 & 2 \end{pmatrix}$

7. $\begin{pmatrix} 2 & -1 & 3 \\ 4 & 1 & 6 \\ 1 & 5 & 3 \end{pmatrix}$ **8.** $\begin{pmatrix} 1 & 0 & 1 & 0 \\ 2 & -1 & 0 & 2 \\ -1 & 0 & 0 & 1 \\ 4 & 1 & -1 & 0 \end{pmatrix}$ **9.** $\begin{pmatrix} a & b & 0 & 0 \\ 0 & a & c & 0 \\ 0 & 0 & a & d \\ 0 & 0 & 0 & a \end{pmatrix}$; $bcd \neq 0$

En los Problemas del 10 *al* 14 *dibuje los círculos de Gershgorin para la matriz dada A y encuentre una cota para* $|\lambda|$ si λ *es un valor característico de A.*

10. $\begin{pmatrix} 2 & 1 & 0 \\ \frac{1}{2} & 5 & \frac{1}{2} \\ 1 & 0 & 6 \end{pmatrix}$ **11.** $\begin{pmatrix} 3 & -\frac{1}{2} & -\frac{1}{3} & 0 \\ 0 & 6 & 1 & 0 \\ \frac{1}{3} & -\frac{1}{3} & 5 & \frac{1}{3} \\ -\frac{1}{2} & \frac{1}{4} & -\frac{1}{4} & 4 \end{pmatrix}$ **12.** $\begin{pmatrix} 1 & 3 & -1 & 4 \\ 2 & 5 & 0 & -7 \\ 3 & -1 & 6 & 1 \\ 0 & 2 & 3 & 4 \end{pmatrix}$

13. $\begin{pmatrix} -7 & \frac{1}{5} & -\frac{1}{5} & \frac{2}{5} \\ -\frac{1}{10} & -10 & \frac{1}{10} & \frac{3}{10} \\ -\frac{1}{4} & \frac{1}{4} & 5 & \frac{1}{4} \\ 0 & -1 & 0 & 4 \end{pmatrix}$ **14.** $\begin{pmatrix} 3 & 0 & -\frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{3} \\ \frac{1}{2} & 5 & -\frac{1}{2} & 0 & 1 & 0 \\ \frac{1}{10} & -\frac{1}{5} & 4 & \frac{3}{5} & -\frac{1}{5} & \frac{1}{10} \\ -1 & 0 & 0 & -3 & 0 & 0 \\ \frac{1}{2} & 0 & -\frac{1}{2} & 0 & 2 & \frac{1}{2} \\ -\frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 & -\frac{1}{4} & 0 \end{pmatrix}$

15. Sea $A = \begin{pmatrix} 2 & \frac{1}{2} & -\frac{1}{3} & \frac{1}{4} \\ \frac{1}{2} & 3 & \frac{1}{2} & 1 \\ -\frac{1}{3} & \frac{1}{2} & 5 & 2 \\ \frac{1}{4} & 1 & 2 & 4 \end{pmatrix}$. Demuestre que los valores característicos de A son números reales positivos.

16. Sea $A = \begin{pmatrix} -4 & 1 & 1 & 1 \\ 1 & -6 & 2 & 1 \\ 1 & 2 & -5 & 1 \\ 1 & 1 & 1 & -4 \end{pmatrix}$. Demuestre que los valores característicos de A son negativos y reales.

17. Sea $P(\lambda) = B_0 + B_1\lambda$ y $Q(\lambda) = C_0 + C_1\lambda$ donde B_0, B_1, C_0 y C_1 son matrices de $n \times n$.
   a. Calcule $F(\lambda) = P(\lambda)Q(\lambda)$.
   b. Sea A una matriz de $n \times n$. Muestre que $F(A) = P(A)Q(A)$ si y sólo si A conmuta con C_0 y C_1.

18. Sea la matriz A de $n \times n$ con valores característicos $\lambda_1, \lambda_2, \ldots, \lambda_n$ y sea $r(A) = \max_{1 \le i \le n} |\lambda_i|$. Si $|A|$ es el máximo de las sumas de las normas de los elementos de los renglones de A definido en la Sección 6.7, muestre que $r(A) \le |A|$.

19. Se dice que la matriz A de $n \times n$ es *estrictamente dominante en la diagonal* si $|a_{ii}| > r_i$ para $i = 1, 2, \ldots, n$, donde r_i está definida por la Ecuación (8). Muestre que si A es estrictamente dominante en la diagonal, entonces $\det A \ne 0$.

Ejercicios de repaso • Capítulo 6

En los Ejercicios del 1 *al* 6 *calcule los valores característicos y espacios característicos de la matriz dada.*

1. $\begin{pmatrix} -8 & 12 \\ -6 & 10 \end{pmatrix}$ **2.** $\begin{pmatrix} 2 & 5 \\ 0 & 2 \end{pmatrix}$ **3.** $\begin{pmatrix} 1 & 0 & 0 \\ 3 & 7 & 0 \\ -2 & 4 & -5 \end{pmatrix}$

4. $\begin{pmatrix} 1 & -1 & 0 \\ 1 & 2 & 1 \\ -2 & 1 & -1 \end{pmatrix}$ **5.** $\begin{pmatrix} 5 & -2 & 0 & 0 \\ 4 & -1 & 0 & 0 \\ 0 & 0 & 3 & -1 \\ 0 & 0 & 2 & 3 \end{pmatrix}$ **6.** $\begin{pmatrix} -2 & 1 & 0 \\ 0 & -2 & 1 \\ 0 & 0 & -2 \end{pmatrix}$

En los Ejercicios del 7 *al* 15 *determine si la matriz A dada es diagonalizable. Si lo es, encuentre una matriz C tal que $C^{-1}AC = D$. Si A es simétrica, encuentre una matriz ortogonal Q tal que $Q^tAQ = D$.*

7. $\begin{pmatrix} -18 & -15 \\ 20 & 17 \end{pmatrix}$ **8.** $\begin{pmatrix} \frac{17}{2} & \frac{9}{2} \\ -15 & -8 \end{pmatrix}$ **9.** $\begin{pmatrix} 1 & 1 & 1 \\ -1 & -1 & 0 \\ -1 & 0 & -1 \end{pmatrix}$

10. $\begin{pmatrix} 4 & 2 & 0 \\ 2 & 4 & 0 \\ 0 & 0 & -3 \end{pmatrix}$ **11.** $\begin{pmatrix} -3 & 2 & 1 \\ -7 & 4 & 2 \\ -5 & 3 & 2 \end{pmatrix}$ **12.** $\begin{pmatrix} 8 & 0 & 12 \\ 0 & -2 & 0 \\ 12 & 0 & -2 \end{pmatrix}$

13. $\begin{pmatrix} 2 & 2 & 0 \\ 2 & 2 & 0 \\ 0 & 0 & -3 \end{pmatrix}$ **14.** $\begin{pmatrix} 4 & 2 & -2 & 2 \\ 1 & 3 & 1 & -1 \\ 0 & 0 & 2 & 0 \\ 1 & 1 & -3 & 5 \end{pmatrix}$ **15.** $\begin{pmatrix} 3 & 4 & -4 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & -4 & 4 & 3 \end{pmatrix}$

En los Ejercicios del 16 *al* 20 *identifique la sección cónica y escríbala en las nuevas variables sin el término xy.*

16. $xy = -4$ **17.** $4x^2 + 2xy + 2y^2 = 8$ **18.** $4x^2 - 3xy + y^2 = 1$
19. $3y^2 - 2xy - 5 = 0$ **20.** $x^2 - 4xy + 4y^2 + 1 = 0$

21. Escriba la forma cuadrática $2x^2 + 4xy + 2y^2 - 3z^2$ en las nuevas variables x', y' y z' de modo que no aparezcan términos cruzados.

En los Ejercicios del 22 *al* 24 *encuentre una matriz C* tal que $C^{-1}AC = J$, *la forma canónica de Jordan de la matriz.*

22. $\begin{pmatrix} -9 & 4 \\ -25 & 11 \end{pmatrix}$ **23.** $\begin{pmatrix} -4 & 4 \\ -1 & 0 \end{pmatrix}$ **24.** $\begin{pmatrix} 0 & -18 & -7 \\ 1 & -12 & -4 \\ -1 & 25 & 9 \end{pmatrix}$

En los Ejercicios del 25 *al* 27 *calcule* e^{At}.

25. $A = \begin{pmatrix} -3 & 4 \\ -2 & 3 \end{pmatrix}$ **26.** $A = \begin{pmatrix} -4 & 4 \\ -1 & 0 \end{pmatrix}$ **27.** $A = \begin{pmatrix} -3 & -4 \\ 2 & 1 \end{pmatrix}$

28. Usando el teorema Cayley-Hamilton, calcule la inversa de $A = \begin{pmatrix} 2 & 3 & 1 \\ -1 & 1 & 0 \\ -2 & -1 & 4 \end{pmatrix}$.

29. Use el teorema del círculo de Gershgorin para encontrar una cota para los valores característicos de

$$A = \begin{pmatrix} 3 & \frac{1}{2} & -\frac{1}{2} & 0 \\ 0 & 4 & \frac{1}{3} & -\frac{1}{3} \\ 1 & 0 & 2 & -1 \\ \frac{1}{2} & -\frac{1}{2} & 1 & -3 \end{pmatrix}$$

CAPÍTULO 7

Métodos numéricos

7.1 El error en los cálculos numéricos

En todos los capítulos de este libro hemos realizado cálculos numéricos. Entre otras cosas, hemos resuelto ecuaciones lineales, multiplicado e invertido matrices, hemos encontrado bases y calculado valores característicos y vectores característicos. Salvo algunas excepciones, los ejemplos se han referido a matrices de 2×2 y 3×3; no porque las aplicaciones prácticas tengan solamente 2 ó 3 variables sino porque de otra manera los cálculos habrían sido muy tediosos.

La situación ha cambiado con el uso difundido y reciente de las calculadoras y las computadoras. Los asombrosos avances de los últimos años en la teoría de métodos numéricos para resolver ciertos problemas, han posibilitado realizar con rapidez y exactitud los cálculos mencionados en el párrafo anterior, con matrices de orden mucho mayor.

Sin embargo, el uso de computadoras presenta nuevas dificultades. Las computadoras no almacenan números como $\frac{2}{3}$, $7\frac{3}{8}$, $\sqrt{2}$, y π. En cambio, todas las computadoras usan lo que se llama *aritmética en punto flotante*. En este sistema, todo número se representa en la forma

$$x = \pm 0.d_1 d_2 \cdots d_k \times 10^n \tag{1}$$

donde $d_1, d_2, \ldots, d_k$ son dígitos individuales enteros positivos y n es entero. Cualquier número escrito en esta forma se denomina *número en punto flotante*. En la Ecuación (1) al número $\pm .d_1 d_2 \cdots d_k$ se le llama *mantisa* y al número n se le denomina *exponente*. El número k se denomina *número de dígitos significativos* de la expresión.

Diferentes computadoras tienen diferentes capacidades en el intervalo de los números que se pueden expresar en la forma de la Ecuación (1). Los dígitos suelen representarse más en forma binaria que en forma decimal. Una computadora de uso común, por ejemplo, utiliza 28 dígitos binarios. Como 2^{28} =

268 435 456, podemos usar los 28 dígitos binarios para representar cualquier número de ocho dígitos, por lo tanto $k = 8$.

Ejemplo 1 Los números siguientes están expresados en la forma de punto flotante:

i. $\frac{1}{4} = 0.25$
ii. $2378 = 0.2378 \times 10^4$
iii. $-0.000816 = -0.816 \times 10^{-3}$
iv. $83.27 = 0.8327 \times 10^2$

Si el número de dígitos significativos fuera ilimitado, no tendríamos problemas. Pero casi siempre que se introducen números en la computadora, los errores se empiezan a acumular. Esto puede suceder de dos maneras:

i. *Truncamiento:* Todos los dígitos significativos después del k-ésimo simplemente son "cortados". Por ejemplo, si se usa truncamiento, $\frac{2}{3} = 0.666666\ldots$ se almacena (con $k = 8$) como $\frac{2}{3} = 0.66666666 \times 10^0$.
ii. *Redondeo:* Si $d_{k+1} \geq 5$, entonces se suma 1 a d_k y el número resultante se trunca. De no ser así, simplemente se trunca el número. Por ejemplo, con redondeo (y $k = 8$), $\frac{2}{3}$ se almacena como $\frac{2}{3} = 0.66666667 \times 10^0$.

Ejemplo 2 Podemos ilustrar cómo se almacenan algunos números con truncamiento y redondeo utilizando ocho dígitos significativos:

Número	Número truncado	Número redondeado
$\frac{8}{3}$	0.26666666×10^1	0.26666667×10^1
π	0.31415926×10^1	0.31415927×10^1
$-\frac{1}{57}$	$-0.17543859 \times 10^{-1}$	$-0.17543860 \times 10^{-1}$

Los errores de redondeo o truncamiento aislados no parecen ser muy significativos. Sin embargo, cuando hay miles de pasos de cómputo, el error *acumulado* puede ser devastador. Entonces, al plantear cualquier esquema numérico, es necesario saber no solamente si se obtendrá la respuesta correcta, teóricamente, sino también cuántos errores de redondeo se acumularán. Para mantener el control, definimos dos tipos de error. Si x es el valor real de un número y x^* es el número que aparece en la computadora, entonces el *error absoluto* ε_a está definido como

$$\varepsilon_a = |x^* - x| \tag{2}$$

En la mayoría de los casos es más interesante el *error relativo* ε_r, que se define como

$$\varepsilon_r = \left|\frac{x^* - x}{x}\right| \tag{3}$$

Ejemplo 3 Sea $x = 2$ y $x^* = 2.1$. Entonces $\varepsilon_a = 0.1$ y $\varepsilon_r = 0.1/2 = 0.05$. Si $x_1 = 2\ 000$ y $x_1^* = 2\ 000.1$, entonces nuevamente $\varepsilon_a = 0.1$. Pero ahora $\varepsilon_r = 0.1/2\ 000 = 0.00005$. La mayoría de las personas estarán de acuerdo en que el error de 0.1 del primer caso es más significativo que el error de 0.1 del segundo.

Gran parte del análisis numérico se refiere a cuestiones de *convergencia* y *estabilidad.* Si x es la solución de un problema y nuestro método de cómputo nos da valores aproximados x_n, entonces el método converge si, teóricamente, x_n se aproxima a x cuando n crece. Si se puede demostrar, además, que los errores de redondeo no se acumulan de modo que hagan la respuesta incierta, entonces el método es estable.

Es fácil dar un ejemplo de un procedimiento en el que el error de redondeo sea muy grande. Supongamos que deseamos calcular $y = 1/(x - 0.66666665)$. Para $x = \frac{2}{3}$, si la computadora trunca, entonces $x = 0.66666666$ y $y = 1/0.00000001 = 10^8 = 10 \times 10^7$. Si la computadora redondea, entonces $x = 0.66666667$ y $y = 1/0.00000002 = 5 \times 10^7$. La diferencia es enorme. La respuesta correcta es $1/(\frac{2}{3} - \frac{66666665}{100000000}) = 60\ 000\ 000 = 6 \times 10^7$.

En este capítulo examinaremos varios esquemas para resolver sistemas de ecuaciones y calcular valores característicos y vectores característicos. Cuando sea posible, también trataremos convergencia y estabilidad. Esto no es más que una vista superficial de métodos numéricos; sin embargo, existen libros y cursos dedicados enteramente a este tema. Para un tratamiento más completo, debería consultar las siguientes obras:

REFERENCIAS SOBRE ÁLGEBRA LINEAL NUMÉRICA

1. Blum, E. K. *Numerical Analysis and Computation: Theory and Practice.* Reading, Mass.: Addison-Wesley, 1972.
2. Burden, R. L., J. D. Faires, y A.C. Reynolds. *Numerical Analysis.* Boston: Prindle, Weber and Schmidt, 1978.
3. Conte, S. D. *Elementary Numerical Analysis, 2a Ed.* Nueva York: McGraw-Hill, 1972.

4. Faddeev, D. K. y V. N. Faddeeva. *Computational Methods of Linear Algebra.* San Francisco: Freeman, 1963.
5. Fox, L. *An Introduction to Numerical Linear Algebra.* New York: Oxford University Press, 1965.

A pesar de que no se incluyen programas de computadora en este capítulo, se dispone de un suplemento especial de este libro (en inglés), que trata de programación en BASIC e incluye varios programas que debería ensayar el lector.

Problemas 7.1

En los Problemas del 1 *al* 13 *convierta el número a un número en punto flotante con ocho decimales de exactitud. Trunque (T) o redondee (R) de acuerdo a lo que se pida.*

1. $\frac{1}{3}$ (T) **2.** $\frac{7}{8}$ **3.** -0.000035 **4.** $\frac{7}{9}$ (R)
5. $\frac{7}{9}$ (T) **6.** $\frac{33}{7}$ (T) **7.** $\frac{85}{11}$ (R) **8.** $-18\frac{5}{6}$ (T)
9. $-18\frac{5}{6}$ (R) **10.** 237,059,628 (T) **11.** 237,059,628 (R)
12. -23.7×10^{15} **13.** 8374.2×10^{-24}

En los Problemas del 14 *al* 21 *se da el número x y una aproximación* x^*. *Encuentre los errores absoluto y relativo* ε_a y ε_r.

14. $x=5;\ x^*=0.49\times 10^1$ **15.** $x=500;\ x^*=0.4999\times 10^3$
16. $x=3720;\ x^*=0.3704\times 10^4$ **17.** $x=\frac{1}{8};\ x^*=0.12\times 10^0$
18. $x=\frac{1}{800};\ x^*=0.12\times 10^{-2}$ **19.** $x=-5\frac{5}{6};\ x^*=-0.583\times 10^1$
20. $x=0.70465;\ x^*=0.70466\times 10^0$
21. $x=70465;\ x^*=0.70466\times 10^5$

7.2 Resolución de sistemas lineales I: Eliminación gaussiana con apoyo

No es difícil programar una computadora para la resolución de un sistema de ecuaciones lineales por el método de eliminación gaussiana o de Gauss-Jordan que se ha usado en este texto. Sin embargo, existe una variante del método, que fue diseñada para reducir el error de redondeo acumulado al resolver un sistema $n \times n$ de ecuaciones. Antes de describir este método recordemos la definición de la forma escalonada de una matriz (Definición 1.6.2).

Definición 1 **Forma de escalonamiento de renglones.** Sea A una matriz de $m \times n$. Entonces A está en *forma escalonada* por renglones si

i. Todos los renglones que tienen sólo ceros aparecen en la parte inferior de la matriz.
ii. El primer número (a partir de la izquierda) de cualquier renglón que no tiene sólo ceros es un 1.
iii. Si dos renglones sucesivos no tienen solamente ceros, el primer 1 del renglón inferior aparece más a la derecha del primer 1 del renglón superior.

Nota. Si A es una matriz triangular superior de $n \times n$ con números 1 en la diagonal principal, entonces es fácil verificar que A está en forma escalonada.

Del Capítulo 1, es evidente que cualquier matriz se puede reducir a la forma escalonada mediante eliminación gaussiana. Sin embargo, hay un problema de cómputo con este método. Si dividimos entre un número pequeño que ha sido redondeado, el resultado puede contener un error de redondeo significativo. Por ejemplo, $1/0.00074 \approx 1351$ mientras que $1/0.0007 \approx 1429$. Para evitar este problema, utilizamos un método llamado *eliminación gaussiana con apoyo* (o *pivoteo*) *parcial*. La idea es dividir siempre entre la componente mayor de una columna, evitando así, en lo posible el tipo de error ilustrado anteriormente. Describiremos el método con un ejemplo simple.

Ejemplo 1 Resuelva el siguiente sistema mediante eliminación gaussiana con apoyo parcial:

$$\begin{aligned} x_1 - x_2 + x_3 &= 1 \\ -3x_1 + 2x_2 - 3x_3 &= -6 \\ 2x_1 - 5x_2 + 4x_3 &= 5 \end{aligned}$$

Solución **Paso 1.** Escriba el sistema en forma de matriz aumentada. De la primera columna con componentes diferentes de cero (llamada *columna de apoyo*) seleccione la componente con el *mayor valor absoluto*. A este elemento se le llama *apoyo* (o *pivote*):

$$\text{apoyo} \longrightarrow \left(\begin{array}{ccc|c} 1 & -1 & 1 & 1 \\ \boxed{-3} & 2 & -3 & -6 \\ 2 & -5 & 4 & 5 \end{array}\right)$$

Paso 2. Ordene los renglones para desplazar el apoyo hasta arriba:

$$\left(\begin{array}{ccc|c} \boxed{-3} & 2 & -3 & -6 \\ 1 & -1 & 1 & 1 \\ 2 & -5 & 4 & 5 \end{array}\right)$$ (se intercambiaron los renglones primero y segundo)

Paso 3. Divida el primer renglón entre el apoyo:

$$\left(\begin{array}{ccc|c} 1 & -\frac{2}{3} & 1 & 2 \\ 1 & -1 & 1 & 1 \\ 2 & -5 & 4 & 5 \end{array}\right)$$ (el primer renglón se divide entre -3)

Paso 4. Sume múltiplos del primer renglón a los otros renglones para hacer que todos los otros elementos de la columna pivote sean iguales a cero:

$$\left(\begin{array}{ccc|c} 1 & -\frac{2}{3} & 1 & 2 \\ 0 & -\frac{1}{3} & 0 & -1 \\ 0 & -\frac{11}{3} & 2 & 1 \end{array}\right)$$ (el primer renglón se multiplica por -1 y -2 y se suma a los renglones segundo y tercero)

Paso 5. Cancele el primer renglón y realice los pasos del 1 al 4 en la *submatriz* resultante:

$$\left(\begin{array}{ccc|c} 1 & -\frac{2}{3} & 1 & 2 \\ 0 & -\frac{1}{3} & 0 & -1 \\ 0 & \boxed{-\frac{11}{3}} & 2 & 1 \end{array}\right)$$

nuevo pivote

$$\left(\begin{array}{ccc|c} 1 & -\frac{2}{3} & 1 & 2 \\ 0 & \boxed{-\frac{11}{3}} & 2 & 1 \\ 0 & -\frac{1}{3} & 0 & -1 \end{array}\right)$$

(se intercambian los renglones primero y segundo de la submatriz)

$$\left(\begin{array}{ccc|c} 1 & -\frac{2}{3} & 1 & 2 \\ 0 & 1 & -\frac{6}{11} & -\frac{3}{11} \\ 0 & -\frac{1}{3} & 0 & -1 \end{array}\right)$$

(el nuevo primer renglón se divide entre el apoyo)

$$\left(\begin{array}{ccc|c} 1 & -\frac{2}{3} & 1 & 2 \\ 0 & 1 & -\frac{6}{11} & -\frac{3}{11} \\ 0 & 0 & -\frac{2}{11} & -\frac{12}{11} \end{array}\right)$$

(el nuevo primer renglón se multiplica por $\frac{1}{3}$ y se suma al nuevo segundo renglón)

Paso 6. Continúe de esta manera hasta que la matriz esté en forma escalonada:

$$\left(\begin{array}{ccc|c} 1 & -\frac{2}{3} & 1 & 2 \\ 0 & 1 & -\frac{6}{11} & -\frac{3}{11} \\ 0 & 0 & \boxed{-\frac{2}{11}} & -\frac{12}{11} \end{array}\right)$$

nuevo pivote

$$\left(\begin{array}{ccc|c} 1 & -\frac{2}{3} & 1 & 2 \\ 0 & 1 & -\frac{6}{11} & -\frac{3}{11} \\ 0 & 0 & 1 & 6 \end{array}\right)$$

(el nuevo primer renglón se divide entre el apoyo)

Paso 7. Haga una *sustitución de regreso* para encontrar la solución (si es que hay alguna) del sistema. Evidentemente, tenemos que $x_3 = 6$. Entonces $x_2 - \frac{6}{11}x_3 = -\frac{3}{11}$ o bien

$$x_2 = -\tfrac{3}{11} + \tfrac{6}{11}x_3 = -\tfrac{3}{11} + \tfrac{6}{11}(6) = 3$$

Finalmente, $x_1 - \frac{2}{3}x_2 + x_3 = 2$ o bien

$$x_1 = 2 + \tfrac{2}{3}x_2 - x_3 = 2 + \tfrac{2}{3}(3) - 6 = -2$$

La solución única está dada por el vector $(-2, 3, 6)$.

Observación. Este ejemplo ilustra el hecho de que es tedioso e ineficiente usar A con mayor valor absoluto, y no simplemente el elemento en la primera columna diferente de cero. El problema con este método es que comúnmente im-

plica una nueva designación de variables cuando se intercambian las columnas para traer el pivote a la primera columna. Por esta razón el método de pivoteo o apoyo parcial descrito anteriormente es más utilizado.

Examinaremos ahora el método aplicado a un sistema más difícil desde el punto de vista de cómputo. Los cálculos se hicieron con una calculadora manual y se redondearon a seis dígitos significativos.

Ejemplo 2 Resuelva el sistema

$$\begin{aligned} 2x_1 - 3.5x_2 + \quad x_3 &= 22.35 \\ -5x_1 + \quad 3x_2 + 3.3x_3 &= -9.08 \\ 12x_1 + 7.8x_2 + 4.6x_3 &= 21.38 \end{aligned}$$

Solución Al usar los pasos descritos anteriormente, obtenemos sucesivamente

$$\left(\begin{array}{ccc|c} 2 & -3.5 & 1 & 22.35 \\ -5 & 3 & 3.3 & -9.08 \\ \boxed{12} & 7.8 & 4.6 & 21.38 \end{array}\right) \xrightarrow{P_{1,3}} \left(\begin{array}{ccc|c} \boxed{12} & 7.8 & 4.6 & 21.38 \\ -5 & 3 & 3.3 & -9.08 \\ 2 & -3.5 & 1 & 22.35 \end{array}\right)$$

pivote

$$\xrightarrow{M_1(\frac{1}{2})} \left(\begin{array}{ccc|c} 1 & 0.65 & 0.383333 & 1.78167 \\ -5 & 3 & 3.3 & -9.08 \\ 2 & -3.5 & 1 & 22.35 \end{array}\right)$$

$$\xrightarrow[A_{1,3}(-2)]{A_{1,2}(5)} \left(\begin{array}{ccc|c} 1 & 0.65 & 0.383333 & 1.78167 \\ 0 & \boxed{6.25} & 5.21667 & -0.17165 \\ 0 & -4.8 & 0.233334 & 18.7867 \end{array}\right)$$

nuevo pivote

$$\xrightarrow{M_2(\frac{1}{6.25})} \left(\begin{array}{ccc|c} 1 & 0.65 & 0.383333 & 1.78167 \\ 0 & 1 & 0.834667 & -0.027464 \\ 0 & -4.8 & 0.233334 & 18.7867 \end{array}\right)$$

$$\xrightarrow{A_{2,3}(4.8)} \left(\begin{array}{ccc|c} 1 & 0.65 & 0.383333 & 1.78167 \\ 0 & 1 & 0.834667 & -0.027464 \\ 0 & 0 & \boxed{4.23974} & 18.6549 \end{array}\right)$$

nuevo pivote

$$\xrightarrow{M_3(\frac{1}{4.23974})} \left(\begin{array}{ccc|c} 1 & 0.65 & 0.383333 & 1.78167 \\ 0 & 1 & 0.834667 & -0.027464 \\ 0 & 0 & 1 & 4.40001 \end{array}\right)$$

La matriz está ahora en forma escalonada por renglones. Al sustituir de regreso obtenemos

$$x_3 = 4.40001$$

$$x_2 = -0.027464 - 0.834667x_3 = -0.027464 - (0.834667)(4.40001)$$

$$= -3.70001$$

$$x_1 = 1.78167-(0.65)(x_2)-(0.383333)x_3 = 1.78167-(0.65)(-3.70001) -(0.383333)(4.40001) = 2.50001$$

La solución correcta es $x_1 = 2.5$, $x_2 = 3.7$ y $x_3 = 4.4$. Nuestras respuestas son, sin duda, muy exactas.

Observación. Este ejemplo ilustra el hecho de que es tedioso e ineficiente usar este método sin una calculadora, especialmente si se requieren varios dígitos significativos de exactitud.

El siguiente ejemplo muestra cómo el pivoteo puede mejorar significativamente las respuestas. Aquí redondeamos a sólo tres dígitos significativos, por lo que se introducen mayores errores de redondeo.

Ejemplo 3 Considere el sistema

$$\begin{aligned} 0.0002x_1-0.00031x_2+0.0017x_3&=0.00609\\ 5x_1 \quad -7x_2 \quad -6x_3&=7\\ 8x_1 \quad +6x_2 \quad +3x_3&=2 \end{aligned}$$

La solución exacta es $x_1=-2$, $x_2=1$, $x_3=4$. Resolvamos el sistema primeramente mediante eliminación gaussiana sin apoyo o pivoteo, redondeando a tres dígitos significativos.

$$\left(\begin{array}{ccc|c} 0.0002 & -0.00031 & 0.0017 & 0.00609\\ 5 & -7 & 6 & 7\\ 8 & 6 & 3 & 2 \end{array}\right) \xrightarrow{M_1(\frac{1}{0.0002})} \left(\begin{array}{ccc|c} 1 & -1.55 & 8.5 & 30.5\\ 5 & -7 & 6 & 7\\ 8 & 6 & 3 & 2 \end{array}\right)$$

$$\xrightarrow[A_{1,3}(-8)]{A_{1,2}(-5)} \left(\begin{array}{ccc|c} 1 & -1.55 & 8.5 & 30.5\\ 0 & 0.75 & -36.5 & -146\\ 0 & 18.4 & -65 & -242 \end{array}\right) \xrightarrow{M_2(\frac{1}{0.75})} \left(\begin{array}{ccc|c} 1 & -1.55 & 8.5 & 30.5\\ 0 & 1 & -48.7 & -195\\ 0 & 18.4 & -65 & -242 \end{array}\right)$$

$$\xrightarrow{A_{2,3}(-18.4)} \left(\begin{array}{ccc|c} 1 & -1.55 & 8.5 & 30.5\\ 0 & 1 & -48.7 & -195\\ 0 & 0 & 831 & 3350 \end{array}\right) \xrightarrow{M_3(\frac{1}{831})} \left(\begin{array}{ccc|c} 1 & -1.55 & 8.5 & 30.5\\ 0 & 1 & -48.7 & -195\\ 0 & 0 & 1 & 4.03 \end{array}\right)$$

Esto produce

$$\begin{aligned} x_3&=4.03\\ x_2&=-195+(48.7)(4.03)=1.26\\ x_1&=30.5+(1.55)(1.26)-8.5(4.03)=-1.8 \end{aligned}$$

Aquí los errores son significativos. Los errores relativos, dados como porcentajes, son

$$x_1:\quad \varepsilon_r=\left|\frac{-0.2}{2}\right|=10\%$$

$$x_2:\quad \varepsilon_r=\left|\frac{0.26}{1}\right|=26\%$$

$$x_3:\quad \varepsilon_r=\left|\frac{0.03}{4}\right|=0.75\%$$

Repitamos ahora el procedimiento *con* pivoteo. Obtenemos (los pivotes aparecen dentro de un círculo)

$$\left(\begin{array}{ccc|c} 0.0002 & -0.00031 & 0.0017 & 0.00609\\ 5 & -7 & 6 & 7\\ \textcircled{8} & 6 & 3 & 2 \end{array}\right) \xrightarrow{P_{1,3}} \left(\begin{array}{ccc|c} \textcircled{8} & 6 & 3 & 2\\ 5 & -7 & 6 & 7\\ 0.0002 & -0.00031 & 0.0017 & 0.00609 \end{array}\right)$$

$$\xrightarrow{M_1(\frac{1}{8})} \left(\begin{array}{ccc|c} 1 & 0.75 & 0.375 & 0.25\\ 5 & -7 & 6 & 7\\ 0.0002 & -0.00031 & 0.0017 & 0.00609 \end{array}\right) \xrightarrow[A_{1,3}(-0.0002)]{A_{1,2}(-5)} \left(\begin{array}{ccc|c} 1 & 0.75 & 0.375 & 0.25\\ 0 & \textcircled{-10.8} & 4.13 & 5.75\\ 0 & -0.00046 & 0.00163 & 0.00604 \end{array}\right)$$

$$\xrightarrow{M_2(\frac{-1}{10.8})} \left(\begin{array}{ccc|c} 1 & 0.75 & 0.375 & 0.25\\ 0 & 1 & -0.382 & -0.532\\ 0 & -0.00046 & 0.00163 & 0.00604 \end{array}\right)$$

$$\xrightarrow{A_{2,3}(0.00046)} \left(\begin{array}{ccc|c} 1 & 0.75 & 0.375 & 0.25\\ 0 & 1 & -0.382 & -0.532\\ 0 & 0 & \textcircled{0.00145} & 0.0058 \end{array}\right)$$

$$\xrightarrow{M_3(\frac{1}{0.00145})} \left(\begin{array}{ccc|c} 1 & 0.75 & 0.375 & 0.25\\ 0 & 1 & -0.382 & -0.532\\ 0 & 0 & 1 & 4.00 \end{array}\right)$$

Por consiguiente,

$$\begin{aligned} x_3&=4.00\\ x_2&=-0.532+(0.382)(4.00)=0.996\\ x_1&=0.25-0.75(0.996)-(0.375)(4.00)=-2.00 \end{aligned}$$

Con pivoteo y tres dígitos significativos de redondeo, x_1 y x_3 se obtienen exactamente y x_2 se obtiene con un error relativo de $0.004/1=0.4$ por ciento.

Antes de terminar esta sección, observemos que hay algunas matrices para las que un error pequeño de redondeo puede tener resultados desastrosos. A tales matrices se las llama *mal condicionadas*.

Ejemplo 4 Considere el sistema

$$\begin{aligned} x_1+\quad\ \ x_2&=1\\ x_1+1.005x_2&=0 \end{aligned}$$

Se ve fácilmente que la solución es $x_1=201$, $x_2=-200$. Si los coeficientes se redondean, con o sin pivoteo, a tres dígitos significativos, obtenemos el sistema

$$\begin{aligned} x_1 + \quad x_2 &= 1 \\ x_1 + 1.01x_2 &= 0 \end{aligned}$$

que tiene solución $x_1=101$, $x_2=-100$. Observe que al redondear se introdujo un error relativo de $0.005/1.005 \approx 0.5$ por ciento en un coeficiente, pero esto indujo un error de ¡casi 50 por ciento en la respuesta final!

Existen técnicas para reconocer y tratar matrices mal condicionadas. Algunas de ellas se tratan en las referencias bibliográficas enumeradas en la primera sección.

Problemas 7.2

En los Problemas del 1 *al* 4 *resuelva el sistema dado mediante eliminación gaussiana con pivoteo parcial. Use una calculadora manual y redondee a seis dígitos significativos en cada paso.*

1. $$\begin{aligned} 2x_1 - x_2 + x_3 &= 0.3 \\ -4x_1 + 3x_2 - 2x_3 &= -1.4 \\ 3x_1 - 8x_2 + 3x_3 &= 0.1 \end{aligned}$$

2. $$\begin{aligned} 4.7x_1 + 1.81x_2 + 2.6x_3 &= -5.047 \\ -3.4x_1 - 0.25x_2 + 1.1x_3 &= 11.495 \\ 12.3x_1 + 0.06x_2 + 0.77x_3 &= 7.9684 \end{aligned}$$

3. $$\begin{aligned} -7.4x_1 + 3.61x_2 + 8.04x_3 &= 25.1499 \\ 12.16x_1 - 2.7x_2 - 0.891x_3 &= 3.2157 \\ -4.12x_1 + 6.63x_2 - 4.38x_3 &= -36.1383 \end{aligned}$$

4. $$\begin{aligned} 4.1x_1 - 0.7x_2 + 8.3x_3 + 3.9x_4 &= -4.22 \\ 2.6x_1 + 8.1x_2 + 0.64x_3 - 0.8x_4 &= 37.452 \\ -5.3x_1 - 0.2x_2 + 7.4x_3 - 0.55x_4 &= -25.73 \\ 0.8x_1 - 1.3x_2 + 3.6x_3 + 1.6x_4 &= -7.7 \end{aligned}$$

En los Problemas 5 *y* 6 *resuelva el sistema mediante eliminación gaussiana con y sin pivoteo, redondeando a tres cifras significativas. Después resuelva el sistema en forma exacta y calcule los errores relativos de los seis valores calculados.*

5. $$\begin{aligned} 0.1x_1 + 0.05x_2 + 0.2x_3 &= 1.3 \\ 12x_1 + 25x_2 - 3x_3 &= 10 \\ -7x_1 + 8x_2 + 15x_3 &= 2 \end{aligned}$$

6. $$\begin{aligned} 0.02x_1 + 0.03x_2 - 0.04x_3 &= -0.04 \\ 16x_1 + 2x_2 + 4x_3 &= 0 \\ 50x_1 + 10x_2 + 8x_3 &= 6 \end{aligned}$$

7. Demuestre que el sistema

$$\begin{aligned} x_1 + \quad x_2 &= 50 \\ x_1 + 1.026x_2 &= 20 \end{aligned}$$

está mal condicionado si se hace el redondeo a tres cifras significativas. ¿Cuál es el error relativo aproximado que se induce por redondeo en cada respuesta?

8. Haga lo mismo para el sistema

$$\begin{aligned} -0.0001x_1 + x_2 &= 2 \\ -x_1 + x_2 &= 3 \end{aligned}$$

7.3 Resolución de sistemas lineales II: Métodos iterativos

En la sección anterior presentamos un método para resolver sistemas lineales *directamente*. Esto es, llevamos a cabo un número fijo de pasos para llegar a una sola respuesta. En análisis numérico este procedimiento es la excepción, más que la regla. Un procedimiento mucho más común es el *iterativo*. En el método iterativo la idea es obtener una *secuencia* de aproximaciones a la solución. Si todo sale bien, esta secuencia converge a la solución correcta, en el sentido de que cada término o *iteración* en la secuencia es una aproximación a la solución, mejor que la que la precede.

Para tener una idea de este método iterativo, considere el siguiente *algoritmo* para calcular $\sqrt{2}$:*

$$x_{n+1} = \frac{1}{2}\left(x_n + \frac{2}{x_n}\right) \qquad (1)$$

Esto significa que empezamos con un valor x_0 y usamos la Ecuación (1) para calcular x_1; a continuación usamos (1) para calcular x_2 y así sucesivamente. En la Tabla 7.1, los cálculos se hicieron en una calculadora manual con 10 dígitos significativos de precisión. Es evidente que el método converge muy rápidamente a la respuesta correcta.

Tabla 7.1

n	x_n	$2/x_n$	$x_n+(2/x_n)$	$x_{n+1}=\frac{1}{2}[x_n+(2/x_n)]$
0	1.0	2.0	3.0	1.5
1	1.5	1.333333333	2.833333333	1.416666667
2	1.416666667	1.411764706	2.828431373	1.414215686
3	1.414215686	1.414211438	2.828427125	1.414213562
4	1.414213562	1.414213562	2.828427125	1.414213562

Existen dos técnicas iterativas comúnmente usadas para resolver un sistema de ecuaciones $A\mathbf{x} = \mathbf{b}$: el *método de Jacobi*† y el *método de Gauss-Seidel*‡. Es-

* El método que se usa aquí, llamado *método de Newton*, fue descubierto en el siglo XVII por Isaac Newton.

† Vea la reseña biográfica que aparece en la página 394.

‡ Encontramos al gran Carl Friedrich Gauss en el Capítulo 1. Así mismo, P.L.V. Seidel (1821-1896) fue otro matemático alemán.

tos métodos se usan bajo ciertas condiciones especiales. Si A está mal condicionada, por ejemplo, entonces como hemos visto, ciertas técnicas directas fallan. Si la matriz A tiene un gran número de ceros (A recibe entonces el nombre de *matriz dispersa*), las técnicas iterativas a menudo producen mejores resultados con menor trabajo. Sin embargo, los dos métodos no siempre convergen. Después de describirlos, examinaremos algunas condiciones bajo las cuales los métodos convergen siempre. En la siguiente discusión supondremos que el det $A \neq 0$, de modo que el sistema tenga una solución única.

I. Método iterativo de Jacobi. Ilustremos el método mediante la solución de un sistema particular. Primero notamos que, debido a que $\det A \neq 0$, A no tiene columnas cero. Entonces, posiblemente mediante un reordenamiento de los renglones de A podemos obtener una nueva matriz de coeficientes A' con elementos diagonales distintos de cero (Problema 14). Entonces suponemos que para la matriz A de $n \times n$, $A = (a_{ij})$, $a_{ii} \neq 0$ para $i = 1, 2, \ldots, n$.

Carl Gustav Jacob Jacobi, 1804-1851



Carl Gustav Jacob Jacobi (*Historical Pictures Service, Chicago*)

Hijo de un próspero banquero, Carl Gustav Jacob Jacobi nació en Potsdam, Alemania, en 1804. Cursó sus estudios en la Universidad de Berlín, en donde recibió su doctorado en 1825. En 1827, fue designado Profesor Extraordinario de Matemáticas en la Universidad de Königsberg. Jacobi dio clases ahí hasta 1842, año en que regresó a Berlín con una pensión del gobierno prusiano. Permaneció en esa ciudad hasta su muerte en 1851.

Un escritor prolífico de tratados matemáticos, Jacobi fue mejor conocido, en su tiempo, por los resultados que obtuvo en la teoría de las funciones elípticas. Actualmente, sin embargo, es mejor recordado por su trabajo en los determinantes. Fue uno de los dos más creativos autores que desarrollaron la teoría de los determinantes; el otro fue Cauchy. En 1829, Jacobi publicó un artículo acerca de álgebra que contenía la notación del jacobiano que utilizamos en la actualidad. En 1841, publicó un extenso tratado con el título *De determinantibus functionalibus*, que estaba dedicado a los resultados del jacobiano. Jacobi puso en relieve la relación entre jacobiano de funciones de varias variables y la derivada de una función de una variable. También demostró que n funciones de n variables son linealmente independientes si y sólo si su jacobiano no es idénticamente igual a cero.

Además de ser un gran matemático, Jacobi fue considerado el mejor maestro de matemáticas de su generación. Inspiró e influyó en un sorprendente número de estudiantes. Para disuadir a sus alumnos de dominar grandes cantidades de matemáticas antes de emprender su propia investigación, Jacobi solía expresar: "El padre de cada uno de ustedes nunca se habría casado, y ustedes no habrían nacido, si él hubiera insistido en conocer a todas las muchachas del mundo antes de casarse con una."

Jacobi creía firmemente en la investigación en las matemáticas puras y frecuentemente la defendía contra la argumentación que sostenía que la investigación debe ser siempre aplicable a algo. Dijo una ocasión: "El verdadero fin de la ciencia es la honra de la mente humana."

Ejemplo 1 Resuelva el sistema

$$\begin{aligned} 4.4x_1 - 2.3x_2 + 0.7x_3 &= -7.43 \\ 0.8x_1 + 2.5x_2 + 1.1x_3 &= 12.17 \\ -1.6x_1 + 0.4x_2 - 5.2x_3 &= 26.12 \end{aligned} \tag{2}$$

Solución Los siguientes cálculos se hicieron con cinco cifras significativas.

Paso 1. Escribamos el sistema (2) de modo que en la i-ésima ecuación, x_i se expresa en términos de las otras variables:

$$\begin{aligned} x_1 &= -\frac{7.43}{4.4} + \frac{2.3}{4.4}x_2 - \frac{0.7}{4.4}x_3 = -1.6886 + 0.52273x_2 - 0.15909\ x_3 \\ x_2 &= \frac{12.17}{2.5} - \frac{0.8}{2.5}x_1 - \frac{1.1}{2.5}x_3 = 4.868 - 0.32x_1 - 0.44x_3 \\ x_3 &= -\frac{26.12}{5.2} - \frac{1.6}{5.2}x_1 + \frac{0.4}{5.2}x_2 = -5.0231 - 0.30769x_1 + 0.076923x_2 \end{aligned} \tag{3}$$

Paso 2. Escogemos arbitrariamente una aproximación inicial de la solución: $x_1^{(0)}, x_2^{(0)}, x_3^{(0)}$. Si no se dispone de otra información, escogemos $x_1^{(0)} = x_2^{(0)} = x_3^{(0)} = 0$.

Paso 3. Sustituimos estos valores iniciales en la parte derecha de (3) para obtener una nueva aproximación $x_1^{(1)}, x_2^{(1)}, x_3^{(1)}$:

$$\begin{aligned} x_1^{(1)} &= -1.6886 + 0 - 0 = -1.6886 \\ x_2^{(1)} &= 4.868 - 0 - 0 = 4.868 \\ x_3^{(1)} &= -5.0231 - 0 + 0 = -5.0231 \end{aligned}$$

Paso 4. Usamos los valores calculados en el Paso 3 para calcular $x_1^{(2)}, x_2^{(2)}, x_3^{(2)}$: y continuar en esta forma para generar las secuencias $\{x_1^{(n)}\}, \{x_2^{(n)}\}, \{x_3^{(n)}\}$.

$$\begin{aligned} x_1^{(2)} &= -1.6886 + 0.52273x_2^{(1)} - 0.15909x_3^{(1)} \\ &= 1.6886 + 0.52273(4.868) - 0.15909(-5.0231) = 1.6552 \\ x_2^{(2)} &= 4.868 - 0.32x_1^{(1)} - 0.44x_3^{(1)} \\ &= 4.868 - 0.32(-1.6886) - 0.44(-5.0231) = 7.6185 \\ x_3^{(2)} &= -5.0231 - 0.30769x_1^{(1)} + 0.076923x_2^{(1)} \\ &= -5.0231 - 0.30769(-1.6886) + 0.076923(4.868) = -4.1291 \end{aligned}$$

Continuamos de esta manera y obtenemos la Tabla 7.2 (redondeada a cinco cifras).

Parece (como podemos ver desde la iteración 8) que las secuencias convergen a los valores $x_1 = 2.5$, $x_2 = 6.4$, $x_3 = -5.3$. Esto se puede verificar por sustitución directa. Note que, al menos en este problema, las iteraciones de Jacobi convergen, aunque más bien lentamente. El método de Gauss-Seidel puede aumentar la rapidez de la convergencia.

Tabla 7.2

Iteración	$x_1^{(n)}$	$x_2^{(n)}$	$x_3^{(n)}$
0	0	0	0
1	−1.6886	4.868	−5.0231
2	1.6552	7.6185	−4.1291
3	2.9507	6.1551	−4.9464
4	2.3158	6.1002	−5.4575
5	2.3684	6.5282	−5.2664
6	2.5617	6.4273	−5.2497
7	2.5063	6.3581	−5.3169
8	2.4808	6.4054	−5.3052
9	2.5037	6.4084	−5.2937
10	2.5034	6.3960	−5.3005
11	2.4980	6.3991	−5.3014
12	2.4998	6.4013	−5.2995
13	2.5006	6.3998	−5.2999
14	2.4999	6.3998	−5.3002
15	2.5000	6.4001	−5.3000

II. Método iterativo de Gauss-Seidel. Si examinamos con cuidado los pasos en el método de Jacobi, notaremos cierta ineficiencia en el Paso 3. Al calcular $x_2^{(n)}$ ya hemos calculado un nuevo valor de $x_1^{(n)}$ pero en vez de éste, hemos usado el valor anterior $x_1^{(n-1)}$. Como las iteraciones están en proceso de convergencia, tiene sentido utilizar la información disponible más reciente.

Ejemplo 2 Resuelva el sistema del Ejemplo 1 utilizando el método de Gauss-Seidel.

Solución Empezamos, como antes, con $x_1^{(0)}=x_2^{(0)}=x_3^{(0)}=0$. Entonces, como anteriormente se señaló

$$x_1^{(1)}=-1.6886+0+0=-1.6886$$

Pero el siguiente paso es diferente. Al utilizar esta nueva aproximación de x_1, obtenemos

$$x_2^{(1)}=4.868-0.32x_1^{(1)}-0.44x_3^{(0)}=4.868-0.32(-1.6886)=5.4084$$

Ahora tenemos nuevas aproximaciones, tanto para x como para x_2. Si utilizamos estos valores en el sistema (3), tenemos que

$$x_3^{(1)}=-5.0231-0.30769x_1^{(1)}+0.076923x_2^{(1)}$$
$$=-5.0231-0.30769(-1.6886)+0.076923(5.4084)=-4.0875$$

Si continuamos de esta manera (siempre utilizando las últimas aproximaciones), obtenemos la Tabla 7.3.

Tabla 7.3

Iteración	$x_1^{(n)}$	$x_2^{(n)}$	$x_3^{(n)}$
0	0	0	0
1	−1.6886	5.4084	−4.0875
2	1.7888	6.0941	−5.1047
3	2.3091	6.3752	−5.2432
4	2.4780	6.3820	−5.2946
5	2.4898	6.4009	−5.2968
6	2.5000	6.3986	−5.3001
7	2.4993	6.4003	−5.2998
8	2.5002	6.3998	−5.3001
9	2.5000	6.4000	−5.3000
10	2.5000	6.4000	−5.3000

Nuevamente nuestro resultado es $x_1=2.5$, $x_2=6.4$, $x_3=-5.3$. Note que las iteraciones de Gauss-Seidel convergen más rápidamente que las iteraciones de Jacobi.

Advertencia. Generalmente (pero no siempre) el método de Gauss-Seidel es más eficiente que el de Jacobi. En el Ejemplo 7 encontramos un sistema para el que las iteraciones del método de Jacobi convergen (lentamente) pero las iteraciones del método de Gauss-Seidel divergen. Lo opuesto es también posible (vea Problema 13).

III. Convergencia. Como se mencionó anteriormente, estos dos métodos no siempre dan una secuencia de iteraciones que converge. A continuación citamos varias condiciones que aseguran la convergencia de las iteraciones. La demostración está fuera del alcance de este texto, sin embargo, para un buen tratamiento del problema consulte el libro de Fox, citado anteriormente.

Definición 1 Sea

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

Entonces A tiene *diagonal estrictamente dominante* si en cada renglón el valor absoluto del elemento diagonal es mayor que la suma de los valores absolutos de los elementos fuera de la diagonal. Esto es:

$$|a_{ii}| > |a_{i1}| + |a_{i2}| + \cdots + |a_{i,i-1}| + |a_{i,i+1}| + \cdots + |a_{in}| = \sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}| \qquad (4)$$

para $i = 1, 2, \ldots, n$

Ejemplo 3 La matriz $A = \begin{pmatrix} 4.4 & -2.3 & 0.7 \\ 0.8 & 2.5 & 1.1 \\ -1.6 & 0.4 & -5.2 \end{pmatrix}$ tiene diagonal estrictamente dominante porque

$$|4.4| > |-2.3| + |0.7| = 3$$
$$|2.5| > |0.8| + |1.1| = 1.9$$

y

$$|-5.2| > |-1.6| + |0.4| = 2$$

Ejemplo 4 Considere el sistema

$$\begin{aligned} x_1 + 3x_2 - x_3 &= 6 \\ 4x_1 - x_2 + x_3 &= 5 \\ x_1 + x_2 - 7x_3 &= -9 \end{aligned} \qquad (5)$$

La matriz $A = \begin{pmatrix} 1 & 3 & -1 \\ 4 & -1 & 1 \\ 1 & 1 & -7 \end{pmatrix}$ *no* tiene diagonal estrictamente dominante. Sin embargo, si intercambiamos las primeras dos ecuaciones del sistema (5) (lo cual, por supuesto, no cambia las soluciones), entonces la matriz del sistema reordenado es $\begin{pmatrix} 4 & -1 & 1 \\ 1 & 3 & -1 \\ 1 & 1 & -7 \end{pmatrix}$, que *sí* es de diagonal estrictamente dominante.

La importancia de que los sistemas tengan matrices de coeficientes con diagonal estrictamente dominante, está dada por el siguiente teorema. El Problema 19 pide su demostración.

Teorema 1 Si A tiene diagonal estrictamente dominante, entonces las iteraciones de ambos métodos, el de Jacobi y el de Gauss-Seidel, convergen a la solución única de $A\mathbf{x} = \mathbf{b}$ para cualquier vector $\mathbf{b}$.

Nota. Dado que la matriz de los Ejemplos 1 y 2 tiene diagonal estrictamente dominante, sabemos antes de hacer cualquier cálculo que ambas secuencias de iteraciones convergen.

Observación. Como veremos en el Ejemplo 6, existen matrices que *no* son de diagonal estrictamente dominante y, sin embargo, ambas secuencias de iteraciones convergen.

Sea A una matriz de $n \times n$. Denotemos como L a la matriz de $n \times n$ que tiene los elementos de A que se encuentran en la parte inferior de la diagonal principal y ceros en los demás lugares; D es la matriz de $n \times n$ con los mismos elementos diagonales de A y cero el resto; U es la matriz con los elementos de A arriba de la diagonal principal y ceros en los demás lugares. A L, D y U se las llama partes *triangular inferior, diagonal* y *triangular superior* de A, respectivamente. Claramente tenemos que

$$A = L + D + U \qquad (6)$$

Ejemplo 5 Sea $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$. Entonces $L = \begin{pmatrix} 0 & 0 & 0 \\ 4 & 0 & 0 \\ 7 & 8 & 0 \end{pmatrix}$, $D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 9 \end{pmatrix}$ y $U = \begin{pmatrix} 0 & 2 & 3 \\ 0 & 0 & 6 \\ 0 & 0 & 0 \end{pmatrix}$.

Teorema 2 Si $r(A)$ denota el valor absoluto del valor característico de A con mayor valor absoluto, entonces, en relación con el sistema $A\mathbf{x} = \mathbf{b}$ con det $A \neq 0$:

i. Las iteraciones de Jacobi convergen si y sólo si

$$r[D^{-1}(L+U)] < 1 \qquad (7)$$

ii. Las iteraciones de Gauss-Seidel convergen si y sólo si

$$r[(D+L)^{-1}U] < 1 \qquad (8)$$

Ejemplo 6 Sea $A = \begin{pmatrix} 2 & 3 \\ 1 & 4 \end{pmatrix}$; entonces $L = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, $D = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}$, $U = \begin{pmatrix} 0 & 3 \\ 0 & 0 \end{pmatrix}$, $D^{-1} = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{4} \end{pmatrix}$, $D^{-1}(L+U) = \begin{pmatrix} 0 & \frac{3}{2} \\ \frac{1}{4} & 0 \end{pmatrix}$, y los valores característicos de $D^{-1}(L+U)$ son $\pm\sqrt{\frac{3}{8}}$. Entonces tenemos que $r[D^{-1}(L+U)] = \sqrt{\frac{3}{8}}$. De manera similar, encontramos que $(D+L)^{-1}U = \begin{pmatrix} 0 & \frac{3}{2} \\ 0 & -\frac{3}{8} \end{pmatrix}$ con valores característicos 0 y $-\frac{3}{8}$. Entonces

$r[(D+L)^{-1}U]=\frac{3}{8}$. Esto nos da un ejemplo de una matriz que no tiene diagonal estrictamente dominante, pero tanto el método de Jacobi como el de Gauss-Seidel convergen.

Ejemplo 7 Sea $A=\begin{pmatrix}1 & 0 & 1\\ -1 & 1 & 0\\ 1 & 2 & -3\end{pmatrix}$. Entonces $L=\begin{pmatrix}0 & 0 & 0\\ -1 & 0 & 0\\ 1 & 2 & 0\end{pmatrix}$, $D=\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & -3\end{pmatrix}$ y $U=\begin{pmatrix}0 & 0 & 1\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}$. Encontramos que $D^{-1}=\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & -\frac{1}{3}\end{pmatrix}$ y $D^{-1}(L+U)=$

$$\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & -\frac{1}{3}\end{pmatrix}\begin{pmatrix}0 & 0 & 1\\ -1 & 0 & 0\\ 1 & 2 & 0\end{pmatrix}=\begin{pmatrix}0 & 0 & 1\\ -1 & 0 & 0\\ -\frac{1}{3} & -\frac{2}{3} & 0\end{pmatrix}$$

La ecuación característica de $D^{-1}(L+U)$ es $\lambda^3+\lambda/3-\frac{2}{3}=0$ con raíces aproximadas $\lambda_1\approx 0.748$, $\lambda_2\approx -0.374+0.868i$, y $\lambda_3\approx -0.374-0.868i$. Tenemos $|\lambda_1|=0.748$, $|\lambda_2|=|\lambda_3|=\sqrt{0.374^2+0.868^2}\approx 0.945$. Entonces $r[D^{-1}(L+U)]\approx 0.945$ y *las iteraciones de Jacobi convergen*. Por otra parte, encontramos que

$$D+L=\begin{pmatrix}1 & 0 & 0\\ -1 & 1 & 0\\ 1 & 2 & -3\end{pmatrix}$$ y

$$(D+L)^{-1}U=\begin{pmatrix}1 & 0 & 0\\ 1 & 1 & 0\\ 1 & \frac{2}{3} & -\frac{1}{3}\end{pmatrix}\begin{pmatrix}0 & 0 & 1\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}=\begin{pmatrix}0 & 0 & 1\\ 0 & 0 & 1\\ 0 & 0 & 1\end{pmatrix}$$

La ecuación característica de $(D+L)^{-1}U$ es $-\lambda^2(\lambda-1)=0$ de modo que $\lambda_1=\lambda_2=0$, $\lambda_3=1$ y $r[(D+L)^{-1}U]=1$. Entonces *las iteraciones de Gauss-Seidel divergen*.

Observación 1. Se puede demostrar además, que la *rapidez* de convergencia en los dos métodos depende de los valores de $r[D^{-1}(L+U)]$ y $r[(D+L)^{-1}U]$. Cuanto menor sea el valor de r, tanto más alta será la rapidez de convergencia.

Observación 2. Denotemos por $|A|$ el máximo de las sumas de las normas de los elementos de los renglones de A (vea Ecuación 6.7.11).

$$|A|=\max_{1\le i\le n}\sum_{j=1}^{n}|a_{ij}| \tag{9}$$

Si se utiliza el teorema del círculo de Gershgorin (vea Problema 18), no es difícil demostrar que para cualquier matriz A de $n\times n$,

$$r(A)\le|A| \tag{10}$$

De modo que el resultado siguiente se deriva directamente del Teorema 2.

Teorema 3 Si A, D, L y U se definen como en el Teorema 2, entonces:

i. La secuencia de iteraciones de Jacobi converge si

$$|D^{-1}(L+U)|<1 \tag{11}$$

ii. La secuencia de iteraciones de Gauss-Seidel converge si

$$|(D+L)^{-1}U|<1 \tag{12}$$

Ejemplo 8 Sea $A=\begin{pmatrix}1 & 0 & \frac{1}{2}\\ \frac{1}{2} & 1 & 1\\ \frac{1}{2} & 0 & 1\end{pmatrix}$. Entonces A no es de diagonal estrictamente dominante, sin embargo aun así podemos demostrar que las iteraciones de Gauss-Seidel convergen. Para $D+L=\begin{pmatrix}1 & 0 & 0\\ \frac{1}{2} & 1 & 0\\ \frac{1}{2} & 0 & 1\end{pmatrix}$, $(D+L)^{-1}=\begin{pmatrix}1 & 0 & 0\\ -\frac{1}{2} & 1 & 0\\ -\frac{1}{2} & 0 & 1\end{pmatrix}$ y

$$(D+L)^{-1}U=\begin{pmatrix}1 & 0 & 0\\ -\frac{1}{2} & 1 & 0\\ -\frac{1}{2} & 0 & 1\end{pmatrix}\begin{pmatrix}0 & 0 & \frac{1}{2}\\ 0 & 0 & 1\\ 0 & 0 & 0\end{pmatrix}=\begin{pmatrix}0 & 0 & \frac{1}{2}\\ 0 & 0 & \frac{3}{4}\\ 0 & 0 & -\frac{1}{4}\end{pmatrix}$$

Entonces $|(D+L)^{-1}U|=\frac{3}{4}<1$.

IV. Análisis del error, o ¿cuándo hay que detenerse? Al resolver problemas mediante métodos iterativos, siempre existe el problema de determinar cuándo parar. Hay dos formas de tomar esta decisión. En la primera, el acuerdo es parar después de un número fijo de iteraciones, digamos 10 ó 20. Sin embargo, como no sabemos cuántas iteraciones se van a requerir para obtener una solución razonablemente exacta, este método no es muy útil.

Una alternativa mejor es parar cuando el error relativo ε_r es suficientemente pequeño. Recordemos que

$$\varepsilon_r=\left|\frac{x^*-x}{x}\right| \tag{13}$$

donde x es la solución exacta y x^* es la aproximación. Por supuesto, no pode-

mos calcular ε_r en forma exacta porque no conocemos la respuesta exacta x (si la conociéramos, no tendríamos ningún problema en primer lugar). Sin embargo, para muchos esquemas numéricos, podemos estimar el error relativo en el esquema de iteración mediante la fórmula

$$\varepsilon_r^{(n)} = \left| \frac{x^{(n)} - x^{(n-1)}}{x^{(n)}} \right| \qquad (14)$$

La Fórmula (14) se puede explicar de la manera siguiente: si sabemos que el esquema converge, entonces la iteración $x^{(n)}$ se acerca más y más a la respuesta "correcta" x. Entonces el error absoluto $\varepsilon_a = |x^{(n)} - x|$ se acerca a cero. Pero entonces, como $x^{(n)} \approx x$, tenemos que $|x^{(n)} - x^{(n-1)}| \approx |x^{(n)} - x|$. Esto significa que la Fórmula (14) se acerca al verdadero error relativo

$$\varepsilon_r = \left| \frac{x^{(n)} - x}{x} \right|$$

Entonces acordamos parar cuando $\varepsilon_r^{(n)}$ es menor que algún valor convencional de ε. Comúnmente $\varepsilon = 0.1$, 0.01, 0.001, o algún valor similar.

Supongamos que en el Ejemplo 1 acordamos iterar hasta que el error relativo estimado $\varepsilon_r^{(n)}$ en el cálculo de x_1 sea menor que 0.01. A partir de la Tabla 7.2 obtenemos la Tabla 7.4

En este caso habríamos parado después de la novena iteración. Esto nos habría dado la estimación $x_1 \approx 2.5037$. Como $x_1 = 2.5$, el error relativo verdadero es $0.0037/2.5 = 0.00148$. Aparentemente, este método solamente nos da una medida burda del error relativo. Sin embargo, es fácil calcular las aproximaciones $\varepsilon_r^{(n)}$. En condiciones de convergencia, nos proporciona una medida razonable de cuánto se acerca uno a la respuesta correcta.

Tabla 7.4

Iteración	$x_1^{(n)}$	$\lvert x_1^{(n)} - x_1^{(n-1)} \rvert$	$\varepsilon_r^{(n)} = \left\lvert \frac{x_1^{(n)} - x_1^{(n-1)}}{x_1^{(n)}} \right\rvert$
0	0		
1	−1.6886	1.6886	1
2	1.6552	3.3438	2.0202
3	2.9507	1.2955	0.43905
4	2.3158	0.6349	0.27416
5	2.3684	0.0526	0.02221
6	2.5617	0.1933	0.07546
7	2.5063	0.0554	0.02210
8	2.4808	0.0255	0.01028
9	2.5037	0.0229	0.00915
10	2.5034	0.0003	0.00012

Problemas 7.3

En los Problemas del 1 al 6 determine si la matriz dada es de diagonal estrictamente dominante.

1. $\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$

2. $\begin{pmatrix} 3 & 3 \\ 4 & 5 \end{pmatrix}$

3. $\begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 1 \end{pmatrix}$

4. $\begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & 1 & -\frac{1}{3} \\ -\frac{1}{2} & -\frac{1}{3} & 1 \end{pmatrix}$

5. $\begin{pmatrix} 3 & -2 & 0 \\ 1 & -4 & 2 \\ -3 & 1 & -5 \end{pmatrix}$

6. $\begin{pmatrix} 6 & -2 & 3 \\ -3 & 5 & 2 \\ -2 & -4 & -7 \end{pmatrix}$

En los Problemas del 7 al 12 resuelva el sistema dado mediante el método de Jacobi y el método de Gauss-Seidel. Realice sus cálculos hasta que el error relativo estimado $\varepsilon_r^{(n)}$ sea menor que el número dado dentro del paréntesis. Empiece con todas las aproximaciones iniciales iguales a cero y utilice cinco dígitos significativos.

7. $\begin{aligned} 2x_1 - x_2 &= 7 \\ 3x_1 + 5x_2 &= 4 \end{aligned}$ (0.01)

8. $\begin{aligned} 3.3x_1 - 2.7x_2 &= -0.6 \\ -4.2x_1 + 8.3x_2 &= 11.95 \end{aligned}$ (0.001)

9. $\begin{aligned} 3x_1 - x_2 + x_3 &= 4 \\ 2x_1 + 5x_2 + 2x_3 &= -5 \\ x_1 + 2x_2 + 4x_3 &= 20 \end{aligned}$ (0.01)

10. $\begin{aligned} 3.8x_1 + 1.6x_2 + 0.9x_3 &= 3.72 \\ -0.7x_1 + 5.4x_2 + 1.6x_3 &= 3.16 \\ 1.5x_1 + 1.1x_2 - 3.2x_3 &= 43.78 \end{aligned}$ (0.001)

11. $\begin{aligned} 5.2x_1 + 3.1x_2 - 1.6x_3 &= 1.64 \\ 1.7x_1 + 2.4x_2 + 0.3x_3 &= 20.42 \\ -6.3x_1 - 3.7x_2 - 12.6x_3 &= 0.27 \end{aligned}$ (0.001)

12. $\begin{aligned} -3.1x_1 + 1.9x_2 - 0.77x_3 &= -12.806 \\ 0.9x_1 - 2.4x_2 + 1.06x_3 &= 12.165 \\ 7.6x_1 - 3.9x_2 + 16.5\,x_3 &= 27.931 \end{aligned}$ (0.0001)

13. Considere el sistema

$$\begin{aligned} x_1 + \tfrac{1}{2}x_2 + \tfrac{1}{2}x_3 &= 2 \\ \tfrac{1}{2}x_1 + x_2 + \tfrac{1}{2}x_3 &= 2 \\ \tfrac{1}{2}x_1 + \tfrac{1}{2}x_2 + x_3 &= 2 \end{aligned}$$

   a. Demuestre que la matriz del sistema no es de diagonal estrictamente dominante.
   b. Empiece con $x_1^{(0)} = x_2^{(0)} = x_3^{(0)} = 0.8$ y demuestre que las iteraciones del método de Jacobi oscilan hacia adelante y hacia atrás, entre los valores 0.8 y 1.2. Esto es, demuestre que diverge la secuencia de iteraciones Jacobi.
   c. Demuestre que las iteraciones de Gauss-Seidel convergen a la solución $x_1 = x_2 = x_3 = 1$, calculando ocho iteraciones y redondeando a cinco cifras significativas.
   ★ d. Explique los resultados de (*b*) y (*c*), a la luz del Teorema 2.

★ 14. Sea A una matriz de $n \times n$ con $\det A \neq 0$. Demuestre que siempre es posible reordenar los renglones de A de manera que los elementos diagonales de A sean todos diferentes de cero.

15. Sea A una matriz diagonal con $\det A \neq 0$. Demuestre que $r[D^{-1}(L + U)] = r[(D + L)^{-1}U] = 0$.

16. Sea A una matriz triangular superior o inferior invertible. Demuestre que las secuencias de iteraciones de los métodos de Jacobi y de Gauss-Seidel siempre convergen.

17. Sea $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Demuestre que las iteraciones tanto del método de Jacobi como las del método de Gauss-Seidel convergen si y sólo si $|bc/ad| < 1$. Esto demuestra que es imposible encontrar un ejemplo de un sistema de 2×2 donde una secuencia de iteraciones converge mientras que la otra no.

★ 18. Use el teorema del círculo de Gershgorin (Teorema 6.8.3) para demostrar que $r(A) \le |A|$, donde $|A| = \max_{1\le i\le n} \sum_{j=1}^{n} |a_{ij}|$.

★ 19. Use la parte (i) del Teorema 3 para demostrar que si A es diagonal estrictamente dominante, entonces las iteraciones del método de Jacobi convergen.

7.4 Cálculo de valores característicos y vectores característicos

Como hemos visto, el cálculo de los valores característicos y de los vectores característicos de una matriz A es importante para una gran variedad de aplicaciones. Para estimar valores característicos, la tendencia natural es encontrar primero el polinomio característico $p(\lambda) = \det(A - \lambda I)$ y después estimar directamente las raíces de $p(\lambda)$. Hay dos problemas con este enfoque; en primer lugar, los polinomios muchas veces están mal condicionados, ésto es, errores pequeños de redondeo en los coeficientes del polinomio pueden conducir a errores grandes en las raíces. El segundo problema es que incluso si los coeficientes de $p(\lambda)$ fueran exactos, es difícil encontrar todas las raíces de un polinomio. Por estas razones, se han diseñado algunas técnicas para calcular valores característicos y vectores característicos directamente. El primero de éstos se utiliza para calcular el valor característico de mayor valor absoluto.

Definición 1 ***Valor característico y vector característico dominantes.*** Sean $\lambda_1, \lambda_2, \ldots, \lambda_n$ los valores característicos de A. Entonces el valor característico λ_1, es *dominante* si

$$|\lambda_1| > |\lambda_i| \quad \text{para } i = 2, \ldots, n \tag{1}$$

Si $\mathbf{v}_1$ es un vector característico de A correspondiente a λ_1, entonces $\mathbf{v}_1$ se denomina *vector característico dominante.*

Ejemplo 1 Si los valores característicos de A son $-4, -2, 1, 3$, entonces -4 es dominante.

Ejemplo 2 Si los valores característicos de A son $-5, 3, 5$, entonces A no tiene valor característico dominante, ya que $|-5| = |5|$.

Ahora describiremos un método, llamado *método de la potencia*, para calcular el valor característico y el vector característico dominantes de una matriz.

I. El método de potencias. Sean $\lambda_1, \lambda_2, \ldots, \lambda_n$ los valores característicos de A y supongamos que

$$|\lambda_1| > |\lambda_2| \ge |\lambda_3| \ge \cdots \ge |\lambda_n| \tag{2}$$

Esto es, λ_1 es el valor característico dominante. Más aún, supongamos que A puede convertirse en una matriz diagonal, o sea, A tiene n vectores característicos linealmente independientes $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$. Sea $\mathbf{x}_0$ un vector en $\mathbb{R}^n$. Existen constantes $c_1, c_2, \ldots, c_n$ tales que

$$\mathbf{x}_0 = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_n\mathbf{u}_n \tag{3}$$

Supóngase que $c_1 \ne 0$. Definimos una secuencia de iteraciones mediante la fórmula

$$\boxed{\mathbf{x}_{n+1} = A\mathbf{x}_n} \tag{4}$$

Entonces

$$\begin{aligned}\mathbf{x}_1 = A\mathbf{x}_0 &= c_1A\mathbf{u}_1 + c_2A\mathbf{u}_2 + \cdots + c_nA\mathbf{u}_n \\ &= c_1\lambda_1\mathbf{u}_1 + c_2\lambda_2\mathbf{u}_2 + \cdots + c_n\lambda_n\mathbf{u}_n\end{aligned}$$

Si continuamos multiplicando por potencias de A resulta que

$$\begin{aligned}\mathbf{x}_2 = A\mathbf{x}_1 = A^2\mathbf{x}_0 = A(A\mathbf{x}_0) &= c_1\lambda_1A\mathbf{u}_1 + c_2\lambda_2A\mathbf{u}_2 + \cdots + c_n\lambda_nA\mathbf{u}_n \\ &= c_1\lambda_1^2\mathbf{u}_1 + c_2\lambda_2^2\mathbf{u}_2 + \cdots + c_n\lambda_n^2\mathbf{u}_n\end{aligned}$$

$$\vdots$$

$$\mathbf{x}_k = A^k\mathbf{x}_0 = c_1\lambda_1^k\mathbf{u}_1 + c_2\lambda_2^k\mathbf{u}_2 + \cdots + c_n\lambda_n^k\mathbf{u}_n \tag{5}$$

o bien

$$\mathbf{x}_k = A^k\mathbf{x}_0 = \lambda_1^k\left[c_1\mathbf{u}_1 + c_2\left(\frac{\lambda_2}{\lambda_1}\right)^k\mathbf{u}_2 + \cdots + c_n\left(\frac{\lambda_n}{\lambda_1}\right)^k\mathbf{u}_n\right] \tag{6}$$

Como $|\lambda_i| < |\lambda_1|$ para $i = 2, 3, \ldots, n$, vemos que $|\lambda_i/\lambda_1|^k$ se acerca a cero a medida que k aumenta. Por lo tanto puede escribirse

$$\boxed{\mathbf{x}_k = A^k\mathbf{x}_0 \approx \lambda_1^k c_1\mathbf{u}_1} \tag{7}$$

Supongamos $\mathbf{u}_1 = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$. Entonces $\lambda_1^k c_1\mathbf{u}_1 = \begin{pmatrix} \lambda_1^k c_1 a_1 \\ \lambda_1^k c_1 a_2 \\ \vdots \\ \lambda_1^k c_1 a_n \end{pmatrix}$.

Ahora supóngase que a_j es diferente de cero. Entonces formamos el cociente

$$\alpha_j^{(k+1)} = \frac{j\text{-ésima componente de } A^{k+1}\mathbf{x}_0}{j\text{-ésima componente de } A^k\mathbf{x}_0} \approx \frac{\lambda_1^{k+1}c_1a_j}{\lambda_1^k c_1 a_j} = \lambda_1 \quad (8)$$

Esto nos da un método para calcular λ_1. Simplemente observamos el cociente de las j ésimas componentes de $\mathbf{x}_{k+1}$ y $\mathbf{x}_k$ y dejamos que k crezca. Más aún, una vez que hemos encontrado λ_1, también conocemos un vector característico correspondiente a λ_1, porque según la Ecuación (7),

$$\mathbf{x}_k \approx \lambda_1^k c_1 \mathbf{u}_1 \quad (9)$$

es un vector característico correspondiente a λ_1, ya que $\lambda_1^k c_1$ es un escalar y $\mathbf{u}_1$ es un vector característico.

Advertencia. El método de potencias que se acaba de describir funciona solamente cuando A tiene un valor característico dominante.*

Ejemplo 3 Use el método de la potencia para encontrar el valor característico y el vector característico dominantes de $A = \begin{pmatrix} -4 & -5 \\ 1 & 2 \end{pmatrix}$.

Solución Aquí $\mathbf{x}_0$ es arbitrario, así que escogemos un valor simple: $\mathbf{x}_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Entonces

$$\mathbf{x}_1 = A\mathbf{x}_0 = \begin{pmatrix} -4 & -5 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -9 \\ 3 \end{pmatrix} \qquad \alpha_1^{(1)} = \tfrac{-9}{1} = -9 \qquad \alpha_2^{(1)} = \tfrac{3}{1} = 3$$

$$\mathbf{x}_2 = A\mathbf{x}_1 = \begin{pmatrix} -4 & -5 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} -9 \\ 3 \end{pmatrix} = \begin{pmatrix} 21 \\ -3 \end{pmatrix} \qquad \alpha_1^{(2)} = \tfrac{21}{-9} = -2.3333 \qquad \alpha_2^{(2)} = \tfrac{-3}{3} = -1$$

Si continuamos de esta manera obtenemos la Tabla 7.5. Todos los resultados están redondeados a cinco cifras significativas.

Parece ser que $\alpha_1^{(k)}$ y $\alpha_2^{(k)}$ convergen a -3, que es el valor característico de A, como puede verificarse fácilmente (el otro es $\lambda_2 = 1$). Mas aún, vemos que $\mathbf{v}_1 = \begin{pmatrix} -49209 \\ 9843 \end{pmatrix}$ es aproximadamente igual a un vector característico de A. Para simplificar este vector, lo "normalizamos" dividiéndolo todo entre su elemento mayor (en valor absoluto) $-49{,}209$ para obtener $\mathbf{v}_1' = \begin{pmatrix} 1 \\ -0.20002 \end{pmatrix} \approx \begin{pmatrix} 1 \\ -\frac{1}{5} \end{pmatrix}$. Se puede verificar fácilmente que $\begin{pmatrix} 1 \\ -\frac{1}{5} \end{pmatrix}$ es un vector característico de A correspondiente al valor característico $\lambda_1 = -3$.

* Se puede demostrar que el método de potencias funciona aun cuando A no puede convertirse en matriz diagonal. Sin embargo, en ese caso la convergencia se logra con una rapidez menor.

Tabla 7.5

Iteración	$\mathbf{x}_k$ (como vector renglón)	$\alpha_1^{(k)}$	$\alpha_2^{(k)}$
0	(1, 1)	—	—
1	(−9, 3)	−9	3
2	(21, −3)	−2.3333	−1
3	(−69, 15)	−3.2857	−5
4	(201, −39)	−2.9130	−2.6
5	(−609, 123)	−3.0299	−3.1538
6	(1821, −363)	−2.9901	−2.9512
7	(−5469, 1095)	−3.0033	−3.0165
8	(16401, −3279)	−2.9989	−2.9945
9	(−49209, 9843)	−3.0004	−3.0018

II. Método de potencias con normalización. En el ejemplo anterior vimos que el tamaño de las iteraciones crece muy rápidamente. Para evitar esto, se realiza lo que hicimos para terminar el problema: *normalizar* o *graduar* el vector $\mathbf{x}_k$ al dividirlo entre su componente mayor (en valor absoluto). Si denotamos como $\mathbf{x}'_k$ a la nueva iteración graduada, entonces

$$\mathbf{x}_{k+1} = A\mathbf{x}'_k \quad (10)$$

Este nuevo método, que se llama *método de potencias con normalización*, origina un vector característico $\mathbf{u}$ cuya mayor componente es 1 y entonces es posible encontrar el valor característico dominante resolviendo la ecuación $A\mathbf{u} = \lambda_1\mathbf{u}$ para λ_1.

Ejemplo 4 Realice nuevamente el Ejemplo 3, utilizando el método de potencias con normalización.

Solución Si $\mathbf{x}_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, entonces $\mathbf{x}_1 = \begin{pmatrix} -9 \\ 3 \end{pmatrix}$ como antes y

$$\mathbf{x}_1' = -\frac{1}{9}\begin{pmatrix} -9 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 \\ -\frac{1}{3} \end{pmatrix}$$

Entonces
$$\mathbf{x}_2 = A\mathbf{x}_1' = \begin{pmatrix} -4 & -5 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ -\frac{1}{3} \end{pmatrix} = \begin{pmatrix} -\frac{7}{3} \\ \frac{1}{3} \end{pmatrix}$$

de modo que

$$\mathbf{x}_2' = -\frac{3}{7}\begin{pmatrix} -\frac{7}{3} \\ \frac{1}{3} \end{pmatrix} = \begin{pmatrix} 1 \\ -\frac{1}{7} \end{pmatrix} = \begin{pmatrix} 1 \\ -0.14286 \end{pmatrix}.$$

Llevamos a cabo más iteraciones en la Tabla 7.6.

Como antes, podemos concluir que $\mathbf{v} = \begin{pmatrix} 1 \\ -0.2 \end{pmatrix}$ es un vector característico de A, correspondiente a λ_1. Así $A\mathbf{v} = \lambda_1\mathbf{v}$ o bien $\begin{pmatrix} -4 & -5 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ -0.2 \end{pmatrix} = \begin{pmatrix} \lambda_1 \\ -0.2\lambda_1 \end{pmatrix}$, que produce $\begin{pmatrix} -3 \\ 0.6 \end{pmatrix} = \begin{pmatrix} \lambda_1 \\ -0.2\lambda_1 \end{pmatrix}$. Entonces $\lambda_1 = -3$.

Tabla 7.6

Iteración	$\mathbf{x}_k$	$\mathbf{x}_k'$ (normalizado)
0	(1, 1)	(1, 1)
1	(−9, 3)	(1, −0.33333)
2	(−2.3333, 0.33333)	(1, −0.14286)
3	(−3.2857, 0.71428)	(1, −0.21739)
4	(−2.9131, 0.56522)	(1, −0.19403)
5	(−3.0299, 0.61194)	(1, −0.20197)
6	(−2.9902, 0.59606)	(1, −0.19934)
7	(−3.0033, 0.60132)	(1, −0.20022)
8	(−2.9989, 0.59956)	(1, −0.19993)
9	(−3.0004, 0.60014)	(1, −0.20002)

Nota. En la Tabla 7.6 la primera componente de $\mathbf{x}_k$ tiende a −3, como debe ser, ya que para k grande,

$$\mathbf{x}_k = A\mathbf{x}_{k-1}' \approx \lambda_1 \mathbf{x}_{k-1}'$$

Pero la primera componente de $\mathbf{x}_{k-1}'$ es 1. De donde la primera componente de $\mathbf{x}_k \approx \lambda_1$. Esto significa que podemos encontrar λ_1 directamente de la tabla.

III. Deflación. El método de potencias tiene la desventaja obvia de que nos da solamente el valor característico dominante. Hay muchas formas de calcular otros valores característicos. Aquí examinamos uno de esos métodos. Primeramente necesitamos el resultado siguiente.

Teorema 1 Consideremos a $\lambda_1, \lambda_2, \ldots, \lambda_n$, los valores característicos de A. Sea λ_1 el valor característico dominante con vector característico $\mathbf{u}_1$, y sea $\mathbf{v}$ un vector columna tal que $\mathbf{u}_1 \cdot \mathbf{v} = 1$. Si la matriz B está dada por

$$B = A - \lambda_1 \mathbf{u}_1 \mathbf{v}^t \qquad (11)$$

entonces los valores característicos de B son $\{0, \lambda_2, \lambda_3, \ldots, \lambda_n\}$.* La demostración de este teorema puede consultarse en la obra de referencia de Blum (p. 239).

Ejemplo 5 En el Ejemplo 3 teníamos que $A = \begin{pmatrix} -4 & -5 \\ 1 & 2 \end{pmatrix}$, $\lambda_1 = -3$, y $\mathbf{u}_1 = \begin{pmatrix} 1 \\ -\frac{1}{5} \end{pmatrix}$. Si $\mathbf{v} = \begin{pmatrix} \frac{1}{2} \\ -\frac{5}{2} \end{pmatrix}$, entonces $\mathbf{u}_1 \cdot \mathbf{v} = 1$ y

$$B = A - \lambda_1 \mathbf{u}_1 \mathbf{v}^t = \begin{pmatrix} -4 & -5 \\ 1 & 2 \end{pmatrix} + 3\begin{pmatrix} 1 \\ -\frac{1}{5} \end{pmatrix}\left(\tfrac{1}{2}, -\tfrac{5}{2}\right)$$

$$= \begin{pmatrix} -4 & -5 \\ 1 & 2 \end{pmatrix} + 3\begin{pmatrix} \frac{1}{2} & -\frac{5}{2} \\ -\frac{1}{10} & \frac{1}{2} \end{pmatrix} = \begin{pmatrix} -4 & -5 \\ 1 & 2 \end{pmatrix} + \begin{pmatrix} \frac{3}{2} & -\frac{15}{2} \\ -\frac{3}{10} & \frac{3}{2} \end{pmatrix} = \begin{pmatrix} -\frac{5}{2} & -\frac{25}{2} \\ \frac{7}{10} & \frac{7}{2} \end{pmatrix}$$

Claramente $\det B = 0$, por tanto, cero es un valor característico de B, tal como se esperaba. El otro valor característico de B es el segundo valor característico de A. Lo calculamos con el método de la potencia con normalización. Empezamos con $x_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, y obtenemos

$$\mathbf{x}_1 = B\mathbf{x}_0 = \begin{pmatrix} -2.5 & -12.5 \\ 0.7 & 3.5 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -15 \\ 4.2 \end{pmatrix}$$

Entonces
$$\mathbf{x}_1' = \begin{pmatrix} 1 \\ -0.28 \end{pmatrix}$$

y
$$\mathbf{x}_2 = B\mathbf{x}_1 = \begin{pmatrix} -2.5 & -12.5 \\ 0.7 & 3.5 \end{pmatrix}\begin{pmatrix} 1 \\ -0.28 \end{pmatrix} = \begin{pmatrix} 1 \\ -0.28 \end{pmatrix}$$

Por lo tanto, sin más problema, vemos que $\begin{pmatrix} 1 \\ -0.28 \end{pmatrix}$ es un vector característico de B, correspondiente al valor característico $\lambda_2 = 1$. Consecuentemente, los valores característicos de $\begin{pmatrix} -4 & -5 \\ 1 & 2 \end{pmatrix}$ son −3 y 1.

* Observe que como $\mathbf{u}_1$ es una matriz de $n \times 1$ (vector columna) y $\mathbf{v}$ también es una matriz de $n \times 1$, entonces $\mathbf{v}^t$ es una matriz de $1 \times n$ y $\mathbf{u}_1\mathbf{v}^t$ es una matriz de $n \times n$.

Nota. No es coincidencia que $B\mathbf{x}_0$ sea un vector característico de B si B es una matriz de 2×2; esto *siempre* se cumple. (Vea el Problema 14, en el que se sugiere por qué es así.)

Ejemplo 6 Calcule los valores característicos de $A = \begin{pmatrix} 4 & -1 & 1 \\ -1 & 3 & -2 \\ 1 & -2 & 3 \end{pmatrix}$ con el método de potencias con normalización y deflación.

Solución Sea $\mathbf{x}_0 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$. Entonces $\mathbf{x}_1 = A\mathbf{x}_0 = \begin{pmatrix} 4 \\ 0 \\ 2 \end{pmatrix}$ y $\mathbf{x}_1' = \begin{pmatrix} 1 \\ 0 \\ 0.5 \end{pmatrix}$. Análogamente $\mathbf{x}_2 = A\mathbf{x}_1 = \begin{pmatrix} 4.5 \\ -2 \\ 2.5 \end{pmatrix}$ y $\mathbf{x}_2' = \begin{pmatrix} 1 \\ -0.44444 \\ 0.55556 \end{pmatrix}$. Continuamos de esta manera y obtenemos los valores en la Tabla 7.7.

Tabla 7.7

Iteración	$\mathbf{x}_k$	$\mathbf{x}_k'$	α_k = primera componente
0	(1, 1, 1)	(1, 1, 1)	1
1	(4, 0, 2)	(1, 0, 0.5)	4
2	(4.5, −2, 2.5)	(1, −0.44444, 0.55556)	4.5
3	(5, −3.4444, 3.5556)	(1, −0.68888, 0.71112)	5
4	(5.4, −4.4889, 4.5111)	(1, −0.83128, 0.83539)	5.4
5	(5.6667, −5.1646, 5.1687)	(1, −0.91139, 0.9[illegible]212)	5.6667
6	(5.8235, −5.5584, 5.5591)	(1, −0.95448, 0.9[illegible]460)	5.8235
7	(5.9091, −5.7726, 5.7728)	(1, −0.97690, 0.9[illegible]693)	5.9091
8	(5.9538, −5.8846, 5.8846)	(1, −0.98838, 0.98838)	5.9538
9	(5.9768, −5.9419, 5.9419)	(1, −0.99416, 0.99416)	5.9768
10	(5.9883, −5.9708, 5.9708)	(1, −0.99708, 0.99708)	5.9883

Resulta que las α_k convergen a $\lambda_1 = 6$ con un vector característico correspondiente $\mathbf{u}_1 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$. Esto se puede verificar fácilmente. A continuación encontramos un vector $\mathbf{v}$ tal que $\mathbf{u}_1 \cdot \mathbf{v} = 1$. Una selección obvia es $\mathbf{v} = \begin{pmatrix} \frac{1}{3} \\ -\frac{1}{3} \\ \frac{1}{3} \end{pmatrix}$.

Entonces

$$\mathbf{u}_1\mathbf{v}^t = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \left(\tfrac{1}{3}, -\tfrac{1}{3}, \tfrac{1}{3}\right) = \begin{pmatrix} \frac{1}{3} & -\frac{1}{3} & \frac{1}{3} \\ -\frac{1}{3} & \frac{1}{3} & -\frac{1}{3} \\ \frac{1}{3} & -\frac{1}{3} & \frac{1}{3} \end{pmatrix}$$

de modo que

$$B = A - \lambda_1\mathbf{u}_1\mathbf{v}^t = \begin{pmatrix} 4 & -1 & 1 \\ -1 & 3 & -2 \\ 1 & -2 & 3 \end{pmatrix} - 6\begin{pmatrix} \frac{1}{3} & -\frac{1}{3} & \frac{1}{3} \\ -\frac{1}{3} & \frac{1}{3} & -\frac{1}{3} \\ \frac{1}{3} & -\frac{1}{3} & \frac{1}{3} \end{pmatrix}$$

$$= \begin{pmatrix} 4 & -1 & 1 \\ -1 & 3 & -2 \\ 1 & -2 & 3 \end{pmatrix} - \begin{pmatrix} 2 & -2 & 2 \\ -2 & 2 & -2 \\ 2 & -2 & 2 \end{pmatrix} = \begin{pmatrix} 2 & 1 & -1 \\ 1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}$$

Vemos que $\det B = 0$; entonces cero es un valor característico de B. Para encontrar el valor característico dominante de B, nuevamente usamos el método de la potencia con normalización. Los resultados se presentan en la Tabla 7.8.

Ahora parece que las iteraciones convergen hacia $\lambda_2 = 3$ y $\mathbf{u}_2 = (1, \frac{1}{2}, -\frac{1}{2})$. Nuevamente, esto puede verificarse fácilmente. A pesar de que $\mathbf{u}_2$ es un vector característico tanto de A como de B, esto no siempre sucede. (En el Ejemplo 5, $\begin{pmatrix} 1 \\ -0.28 \end{pmatrix}$ era un vector característico de B pero no de A.)

Tabla 7.8

Iteración	$\mathbf{x}_k$	$\mathbf{x}_k'$	α_k = primera componente de $\mathbf{x}_k$
0	(1, 1, 1)	(1, 1, 1)	1
1	(2, 2, 0)	(1, 1, 0)	2
2	(3, 2, −1)	(1, 0.66667, −0.33333)	3
3	(3, 1.6667, −1.3333)	(1, 0.55556, −0.44443)	3
4	(3, 1.5556, −1.4444)	(1, 0.51853, −0.48147)	3
5	(3, 1.5185, −1.4815)	(1, 0.50617, −0.49383)	3
6	(3, 1.5062, −1.4938)	(1, 0.50207, −0.49793)	3
7	(3, 1.5021, −1.4979)	(1, 0.50070, −0.49930)	3

Finalmente, se utilizará la deflación otra vez para encontrar el último valor característico de B (y por lo tanto de A). Hacemos $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$. Entonces $\mathbf{u}_2 \cdot \mathbf{v}_1 = 1$ y $\mathbf{u}_2\mathbf{v}_1^t = \begin{pmatrix} 1 \\ \frac{1}{2} \\ -\frac{1}{2} \end{pmatrix} (1, 0, 0) = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 0 & 0 \\ -\frac{1}{2} & 0 & 0 \end{pmatrix}$ y $C = B - \lambda_2\mathbf{u}_2\mathbf{v}_1^t = \begin{pmatrix} 2 & 1 & -1 \\ 1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} - \begin{pmatrix} 3 & 0 & 0 \\ \frac{3}{2} & 0 & 0 \\ -\frac{3}{2} & 0 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 1 & -1 \\ -\frac{1}{2} & 1 & 0 \\ \frac{1}{2} & 0 & 1 \end{pmatrix}$. Omitimos la iteración, que demuestra que el valor característico dominante de C es $\lambda_3 = 1$. De modo que los valores característicos de A son 6, 3 y 1.

El método de potencias, junto con deflación, nos da una forma razonable de encontrar los valores característicos de A, si ésta no tiene dos valores característicos con el mismo valor absoluto y si es buena la aproximación de cada valor característico. Por ejemplo, si λ_1 es inexacta, el cálculo de λ_2 mediante deflación puede ser bastante más inexacta.

Hay muchas otras formas de calcular numéricamente los valores característicos de una matriz cuadrada. Un método que funciona bien con una matriz simétrica es el llamado *método de Jacobi*. La idea es calcular una secuencia de matrices ortogonales cuyos elementos diagonales se aproximen a los valores característicos de A. Muchas de las referencias mencionadas en la Sección 7.1 tratan este método. Finalmente, notamos que la decisión de "cuándo parar" puede tomarse, como en la última sección, calculando valores aproximados del error relativo $\varepsilon_r^{(n)}$.

Problemas 7.4

En los Problemas del 1 al 6 calcule el valor característico dominante y el vector característico de A mediante el método de potencias con normalización.

1. $\begin{pmatrix} -2 & -2 \\ -5 & 1 \end{pmatrix}$
2. $\begin{pmatrix} 8 & 3 \\ -3 & -2 \end{pmatrix}$
3. $\begin{pmatrix} -22.3 & -32 \\ 12 & 17.7 \end{pmatrix}$
4. $\begin{pmatrix} 1 & -1 & 4 \\ 3 & 2 & -1 \\ 2 & 1 & -1 \end{pmatrix}$
5. $\begin{pmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{pmatrix}$
6. $\begin{pmatrix} 5 & 4 & 2 \\ 4 & 5 & 2 \\ 2 & 2 & 2 \end{pmatrix}$
7. Use el método de potencias para estimar el valor característico dominante de $A = \begin{pmatrix} 1 & 7 \\ 6 & 3 \end{pmatrix}$:
   a. Redondee a cinco cifras significativas y continúe las iteraciones hasta que el error relativo estimado $\varepsilon_r^{(n)} < 0.001$.
   b. Calcule el valor característico dominante en forma exacta. ¿Cuál es el valor exacto de ε_r?
8. Para la matriz $A = \begin{pmatrix} -16.32 & 13 \\ 8 & 4.79 \end{pmatrix}$, siga los pasos del Problema 7. Use seis cifras significativas.
9. Demuestre que las iteraciones del método de potencias no convergen para la matriz $A = \begin{pmatrix} -3 & 5 \\ -2 & 3 \end{pmatrix}$. Explique por qué.
10. Haga lo mismo para la matriz $A = \begin{pmatrix} 2 & -1 \\ 5 & -2 \end{pmatrix}$.

En los Problemas del 11 al 13 use deflación para encontrar los otros valores característicos.

11. Para la matriz del Problema 1
12. Para la matriz del Problema 3
13. Para la matriz del Problema 4

14. Sea $A = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$. Demuestre que para cualquier vector de dos elementos $\mathbf{x}_0$, se tiene que $B\mathbf{x}_0$ es un vector característico de B, donde B está definida en la Ecuación (11).

Ejercicios de repaso • Capítulo 7

En los Ejercicios del 1 al 4 el número x y una aproximación x son conocidos. Encuentre los errores absoluto y relativo ε_a y ε_r.*

1. $x = -7$; $x^* = -6.98$
2. $x = 1000$; $x^* = 1.002 \times 10^3$
3. $x = \frac{1}{75}$; $x^* = 1 \times 10^{-2}$
4. $x = 37539$; $x^* = 3.7 \times 10^4$
5. Reduzca la matriz $\begin{pmatrix} 2 & -4 & 6 \\ 1 & -3 & 5 \\ -4 & 9 & -13 \end{pmatrix}$ a la forma escalonada por renglón.

En los Ejercicios 6 y 7 resuelva el sistema dado mediante la eliminación gaussiana con apoyo o pivoteo parcial. Redondee a seis dígitos significativos en cada paso.

6. $$\begin{aligned} 3.6x_1 + 8.2x_2 - 6.4x_3 &= 1.26 \\ -4.5x_1 - 5.9x_2 + 0.3x_3 &= 2.57 \\ 0.7x_1 + 3.6x_2 - 4.8x_3 &= 2.15 \end{aligned}$$
7. $$\begin{aligned} 1.3x_1 - 9.6x_2 + 5.35x_3 &= 0.515 \\ -12x_1 - 15x_2 + 3.8x_3 &= -71.966 \\ 1.06x_1 - 22.2x_2 + 9.93x_3 &= 1.809 \end{aligned}$$

En los Ejercicios 8 y 9 determine si la matriz dada es de diagonal estrictamente dominante.

8. $\begin{pmatrix} -1 & \frac{1}{2} & \frac{1}{3} \\ -\frac{5}{6} & 1 & 0 \\ 2 & \frac{3}{2} & -4 \end{pmatrix}$
9. $\begin{pmatrix} -1 & \frac{1}{3} & -\frac{1}{3} \\ -\frac{5}{6} & 1 & \frac{1}{6} \\ 2 & \frac{3}{2} & -4 \end{pmatrix}$

En los Ejercicios 10 y 11 resuelva el sistema dado mediante los métodos de Jacobi y de Gauss-Seidel. Lleve a cabo las iteraciones hasta que el error relativo estimado $\varepsilon_r^{(n)}$ sea menor que el número dado entre paréntesis. Use seis cifras significativas en todos los cálculos.

10. $$\begin{aligned} 2.7x_1 - 0.9x_2 + 1.3x_3 &= 6.98 \\ -0.3x_1 + x_2 + 0.4x_3 &= -2.77 \quad (0.001) \\ 4x_1 - 3.3x_2 + 9.6x_3 &= 21.79 \end{aligned}$$
11. $$\begin{aligned} 42.31x_1 + 8.62x_2 + 19.4x_3 &= -2.2502 \\ -4.73x_1 + 80.4x_2 - 37.2x_3 &= 3.5402 \quad (0.0001) \\ 8.37x_1 + 30.9x_2 - 57.4x_3 &= -24.0858 \end{aligned}$$

En los Ejercicios del 12 al 14 estime el valor característico dominante y el vector característico mediante el método de la potencia con normalización.

12. $\begin{pmatrix} 8 & -2 \\ 4 & 2 \end{pmatrix}$
13. $\begin{pmatrix} -6 & 3 \\ 6 & 1 \end{pmatrix}$
14. $\begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix}$
15. Use deflación para encontrar el segundo valor característico de la matriz del Ejercicio 13.

APÉNDICE 1

Inducción matemática

La *inducción matemática*‡ es el nombre sofisticado que se le ha dado a un principio lógico simple que puede usarse para probar afirmaciones matemáticas de cierto tipo. Típicamente se usa la inducción matemática para demostrar que una cierta proposición o una cierta ecuación son válidas para todo entero positivo. Por ejemplo, podría requerirse demostrar que $2^n > n$ para todos los enteros $n \geq 1$.

Para lograr esto, se procede en dos pasos.

> **i.** Probamos que la proposición es verdadera para algún entero N (usualmente $N = 1$).
>
> **ii.** Se *supone* que la proposición es verdadera para un entero k, y se *demuestra* que es verdadera para $k + 1$.

Si podemos completar estos dos pasos, habremos demostrado entonces la validez de la proposición en cuestión para *todos* los enteros mayores que o iguales a N. Para convencerse de este hecho, razonemos como sigue: como la proposición es verdadera para N (por el paso *i*) también será válida para el entero $N + 1$ (por el paso *ii*). Entonces es también cierta para el entero $(N + 1) + 1 = N + 2$ (de nuevo, por el paso *ii*), y así sucesivamente. Se demostrará el procedimiento con algunos ejemplos.

Ejemplo 1 Demostrar que $2^n > n$ para todos los enteros $n \geq 1$.

Solución **i.** Si $n = 1$, entonces $2^n = 2 = 2 > 1 = n$, así que $2^n > n$ cuando $n = 1$.

ii. Supóngase que $2^k > k$ cuando $k > 1$ es un entero.

‡ Esta técnica fue utilizada por primera vez en una demostración matemática por el gran matemático francés Pierre de Fermat (1601-1665).

Entonces

$$2^{k+1} = 2 \cdot 2^k = 2^k + 2^k \overset{\text{ya que } 2^k > k}{>} k + k > k + 1$$

Esto completa la demostración, ya que hemos demostrado que $2^1 > 1$, lo cual implica, por el paso (*ii*), que $2^2 > 2$ y, por el paso (*ii*) de nuevo, que $2^3 > 3$, $2^4 > 4$, y así sucesivamente.

Ejemplo 2 Usar la inducción matemática para demostrar la fórmula para la adición de los primeros n enteros:

$$1 + 2 + 3 + \cdots + n = \frac{n(n+1)}{2} \quad (1)$$

Solución

i. Si $n = 1$, la suma del primer entero es 1. Pero $(1)(1 + 2)/2 = 1$, así que la Ecuación (1) se satisface en el caso en que $n = 1$.

ii. Supóngase que (1) se verifica para $n = k$, esto es:

$$1 + 2 + 3 + \cdots + k = \frac{k(k+1)}{2} \quad (2)$$

Debe demostrarse que se cumple para $n = k + 1$. Esto es, hay que probar que

$$1 + 2 + 3 + \cdots + k + (k+1) = \frac{(k+1)(k+2)}{2}$$

Pero

$$1 + 2 + 3 + \cdots + k + (k+1) = (1 + 2 + 3 + \cdots + k) + (k+1)$$

$$\overset{\text{por (2)}}{=} \frac{k(k+1)}{2} + (k+1)$$

$$= \frac{k(k+1) + 2(k+1)}{2}$$

$$= \frac{(k+1)(k+2)}{2}$$

y la demostración queda completa. El lector puede intentar resolver varios ejemplos para cerciorarse de que la fórmula (1) funciona realmente. Por ejemplo

$$1 + 2 + 3 + 4 + 5 + 6 + 7 + 8 + 9 + 10 = \frac{10(11)}{2} = 55$$

Ejemplo 3 Usar inducción matemática para probar la fórmula para la adición de los cuadrados de los primeros n enteros positivos:

$$1^2 + 2^2 + 3^2 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6} \quad (3)$$

Solución

i. Como $1(1 + 1)(2 \cdot 1 + 1)/6 = 1 = 1^2$, la Ecuación (3) es válida cuando $n = 1$.

ii. Supóngase que la Ecuación (3) se verifica para $n = k$, esto es,

$$1^2 + 2^2 + 3^2 + \cdots + k^2 = \frac{k(k+1)(2k+1)}{6} \quad (4)$$

Entonces, para probar que (3) es cierta para $n = k + 1$, tenemos

$$1^2 + 2^2 + 3^2 + \cdots + k^2 + (k+1)^2 \overset{\text{por (4)}}{=} \frac{k(k+1)(2k+1)}{6} + (k+1)^2$$

$$= \frac{k(k+1)(2k+1) + 6(k+1)^2}{6}$$

$$= \frac{k+1}{6}[k(2k+1) + 6(k+1)]$$

$$= \frac{k+1}{6}[2k^2 + 7k + 6]$$

$$= \frac{k+1}{6}[(k+2)(2k+3)]$$

$$= \frac{(k+1)(k+2)[2(k+1)+1]}{6}$$

que es la Ecuación (3) para $n = k + 1$ y la demostración queda terminada. De nuevo, el lector quizá desee experimentar con esta fórmula. Por ejemplo

$$1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2 + 7^2 = \frac{7(7+1)(2 \cdot 7 + 1)}{6}$$

$$= \frac{7 \cdot 8 \cdot 15}{6} = 140$$

Ejemplo 4 Si $a \neq 1$, utilizar inducción matemática para demostrar la fórmula de la suma de una progresión geométrica:

$$1 + a + a^2 + \cdots + a^n = \frac{1 - a^{n+1}}{1 - a} \quad (5)$$

Solución

i. Si $n = 0$, entonces

$$\frac{1 - a^{0+1}}{1 - a} = \frac{1 - a}{1 - a} = 1$$

Así que la Ecuación (5) es válida si $n = 0$. (Usamos $n = 0$ en vez de $n = 1$ puesto que $a^0 = 1$ es el primer término.)

ii. Supóngase que (5) se verifica para $n = k$; esto es,

$$1 + a + a^2 + \cdots + a^k = \frac{1 - a^{k+1}}{1 - a} \tag{6}$$

Entonces,

$$1 + a + a^2 + \cdots + a^k + a^{k+1} \overset{\text{por (6)}}{=} \frac{1 - a^{k+1}}{1 - a} + a^{k+1}$$

$$= \frac{1 - a^{k+1} + (1 - a)a^{k+1}}{1 - a} = \frac{1 - a^{k+2}}{1 - a}$$

de manera que la Ecuación (6) también es válida cuando $n = k + 1$ y la demostración se completa.

Ejemplo 5 Sean $A_1, A_2, \ldots, A_k$, k matrices de $n \times n$ invertibles. Demostrar que

$$(A_1A_2 \cdots A_m)^{-1} = A_m^{-1}A_{m-1}^{-1} \cdots A_2^{-1}A_1^{-1} \tag{7}$$

Si $m = 2$, tenemos, por el Teorema 1.8.3, que $(A_1A_2)^{-1} = A_2^{-1}A_1^{-1}$. Por lo tanto, la Ecuación (7) se verifica para $m = 2$. Supondremos que es cierto si $m = k$ y se la demostrará para $m = k + 1$. Sea $B = A_1A_2 \cdots A_k$. Entonces

$$(A_1A_2 \cdots A_kA_{k+1})^{-1} = (BA_{k+1})^{-1} = A_{k+1}^{-1}B^{-1} \tag{8}$$

Pero, por la hipótesis de inducción,

$$B^{-1} = (A_1A_2 \cdots A_k)^{-1} = A_k^{-1}A_{k-1}^{-1} \cdots A_2^{-1}A_1^{-1} \tag{9}$$

Sustituyendo (9) en (8) se concluye la demostración.

Problemas A.1

En los siguientes problemas demuestre el resultado requerido usando inducción matemática.

1. Muestre que $1^3 + 2^3 + 3^3 + \cdots + n^3 = [n^2(n+1)^2]/4$.
2. Sean $\mathbf{u}, \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$, $n + 1$ vectores en $\mathbb{R}^2$. Demostrar que
$$\mathbf{u} \cdot (\mathbf{v}_1 + \mathbf{v}_2 + \cdots + \mathbf{v}_n) = \mathbf{u} \cdot \mathbf{v}_1 + \mathbf{u} \cdot \mathbf{v}_2 + \cdots + \mathbf{u} \cdot \mathbf{v}_n.$$
3. Demostrar que si $a \neq 1$,
$$1 + 2a + 3a^2 + \cdots + na^{n-1} = \frac{1 - (n+1)a^n + na^{n+1}}{(1-a)^2}$$
4. Demuestre que si un conjunto S contiene n elementos entonces S tiene 2^n subconjuntos.
5. Suponiendo que todo polinomio tiene al menos una raíz, demuestre que un polinomio de grado n tiene exactamente n raíces (contando las multiplicidades).
6. Dado que $\det AB = \det A \det B$ demuestre que $\det A_1A_2 \ldots A_m = \det A_1 \det A_2 \ldots \det A_m$, donde $A_1, \ldots, A_m$ son matrices de $n \times n$.
7. Si $A_1, A_2, \ldots, A_k$ son matrices de $m \times n$ muestre que $(A_1 + A_2 + \cdots + A_k)^t = A_1^t + A_2^t + \cdots + A_k^t$. Se puede suponer que $(A + B)^t = A^t + B^t$.

APÉNDICE **2**

Números complejos

En el Capítulo 6 nos encontramos con el problema de hallar las raíces del polinomio

$$\lambda^2 + a\lambda + b = 0 \tag{1}$$

Para encontrar las raíces usamos la fórmula cuadrática

$$\lambda = \frac{-a \pm \sqrt{a^2 - 4b}}{2} \tag{2}$$

Si $a^2 - 4b > 0$ existen dos raíces reales. Si $a^2 - 4b = 0$ obtenemos la raíz simple (de multiplicidad 2) $\lambda = -a/2$. Para tratar el caso $a^2 - 4b < 0$ introducimos el *número imaginario**

$$i = \sqrt{-1} \tag{3}$$

* El lector no debe preocuparse con el término "imaginario". Es sólo un nombre. El matemático británico Alfred North Whitehead, en el capítulo sobre números imaginarios de su *Introduction to Mathematics*, escribió lo siguiente:

Quizás sea útil observar aquí que cierto tipo de mentalidad siempre se preocupa y preocupa a otros examinando la aplicabilidad de los términos técnicos. ¿Es correcto llamar números a los números inconmensurables?, ¿son realmente números los números positivos y negativos?, ¿son los números imaginarios realmente imaginarios, y son números?, con preguntas ociosas de ese tipo. Es necesario comprender muy claramente que, en ciencia, los términos técnicos son nombres asignados arbitrariamente, como los nombres de pila que se ponen a los niños. No se plantea la cuestión de si los nombres son correctos o erróneos. Pueden ser apropiados o inapropiados, porque a veces pueden elegirse de manera que sean fáciles de recordar, o sugieren ideas pertinentes e importantes. Pero el principio esencial es el que Humpty Dumpty explicó muy claramente a Alicia en el País de las Maravillas, cuando le dijo, respecto del uso de las palabras: "Les pago extra y las hago significar lo que quiero". De manera que no nos preocuparemos de si los números imaginarios son imaginarios, o de si son realmente números, sino que consideraremos el término como un nombre arbitrario de una determinada idea matemática, que trataremos ahora de explicar.

Entonces, para $a^2-4b<0$,

$$\sqrt{a^2-4b}=\sqrt{(4b-a^2)(-1)}=\sqrt{4b-a^2}\,i$$

y las dos raíces de (1) están dadas por

$$\lambda_1=-\frac{a}{2}+\frac{\sqrt{4b-a^2}}{2}i \quad \text{y} \quad \lambda_2=-\frac{a}{2}-\frac{\sqrt{4b-a^2}}{2}i$$

Ejemplo 1 Encuentre las raíces de la ecuación cuadrática $\lambda^2+2\lambda+5=0$.

Solución Tenemos $a=2$, $b=5$ y $a^2-4b=-16$. Así, $\sqrt{a^2-4b}=\sqrt{-16}=\sqrt{16}\sqrt{-1}=4i$ y las raíces son

$$\lambda_1=\frac{-2+4i}{2}=-1+2i \quad \text{y} \quad \lambda_2=-1-2i$$

Definición 1 Un *número complejo* es un número de la forma

$$z=\alpha+i\beta \tag{4}$$

donde α y β son números reales. α se conoce como la *parte real* de z y se denota Re z. β se conoce como *parte imaginaria* de z y se denota Im z. La representación (4) se conoce como *forma cartesiana* del número complejo z.

Observación. Si $\beta=0$ en la Ecuación (4) entonces $z=\alpha$ es un número real. En este contexto podemos considerar al conjunto de números reales como un subconjunto del conjunto de números complejos.

Ejemplo 2 En el Ejemplo 1 Re $\lambda_1=-1$, Im $\lambda_1=2$.

Las reglas ordinarias de la suma y la multiplicación se cumplen para los números complejos.

Ejemplo 3 Sean $z=2+3i$ y $w=5-4i$. Calcule **(i)** $z+w$, **(ii)** $3w-5z$ y **(iii)** zw.

Solución

i. $z+w=(2+3i)+(5-4i)=(2+5)+(3-4)i=7-i$.
ii. $3w=3(5-4i)=15-12i$; $5z=10+15i$; y $3w-5z=(15-12i)-(10+15i)=(15-10)+i(-12-15)=5-27i$.
iii. $zw=(2+3i)(5-4i)=(2)(5)+2(-4i)+(3i)(5)+(3i)(-4i)=10-8i+15i-12i^2=10+7i+12=22+7i$. Aquí hemos usado el hecho de que $i^2=-1$.

Los números complejos se pueden representar como puntos en el *plano complejo*, con Re z representada a lo largo del eje x e Im z a lo largo del eje y. Algunos puntos representativos se muestran en la Figura A.1.

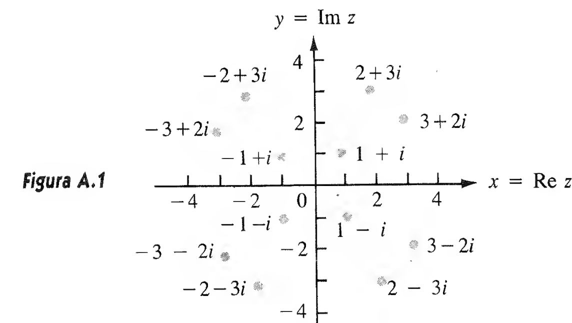


Figura A.1

Si $z=\alpha+i\beta$, entonces definimos el *conjugado* de z, denotado como $\bar{z}$, por

$$\bar{z}=\alpha-i\beta \tag{5}$$

En la Figura A.2 se muestra un valor representativo de z y $\bar{z}$.

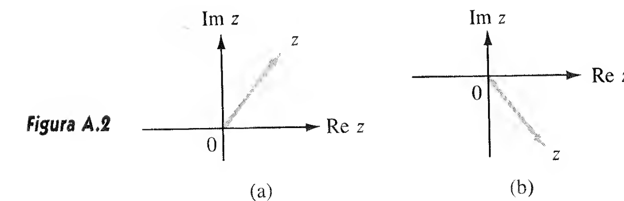


Figura A.2

Ejemplo 4 Calcule el conjugado de **(i)** $1+i$, **(ii)** $3-4i$, **(iii)** $-7+5i$, y **(iv)** -3.

Solución **(i)** $\overline{1+i}=1-i$; **(ii)** $\overline{3-4i}=3+4i$; **(iii)** $\overline{-7+5i}=-7-5i$; **(iv)** $\overline{-3}=-3$.

No es difícil mostrar (Problema 35) que

$$\bar{z}=z \quad \text{si y sólo si } z \text{ es real} \tag{6}$$

Si $z=\beta i$ con β real, entonces z se conoce como *imaginario puro*. Podemos entonces mostrar (Problema 36) que

$$\bar{z}=-z \quad \text{si y sólo si } z \text{ es imaginario puro} \tag{7}$$

Sea $p_n(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$ un polinomio con coeficientes reales. Entonces se puede mostrar (Problema 41) que las raíces complejas de la ecuación $p_n(x) = 0$ ocurren en pares de conjugados complejos. Esto es, si z es una raíz de $p_n(x) = 0$, entonces $\bar{z}$ también lo es. Vimos este hecho ilustrado en el Ejemplo 1 en el caso $n = 2$.

Para $z = \alpha + i\beta$, definimos la *magnitud* de z, denotada como $|z|$, por

$$|z| = \sqrt{\alpha^2 + \beta^2} \tag{8}$$

y definimos el *argumento* de z, denotado arg z, como el ángulo θ entre la recta $\overline{0z}$ y la parte positiva del eje x. De la Figura A.3 vemos que $r = |z|$ es la distancia de z al origen y

$$\theta = \arg z = \tan^{-1}\frac{\beta}{\alpha} \tag{9}$$

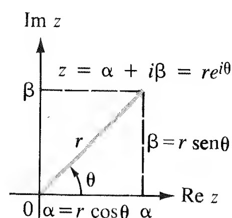


Figura A.3

Por convención siempre escogemos el valor de $\tan^{-1}\beta/\alpha$ que está en el intervalo

$$-\pi < \theta \leq \pi \tag{10}$$

De la Figura A.4 vemos que

$$|\bar{z}| = |z| \tag{11}$$

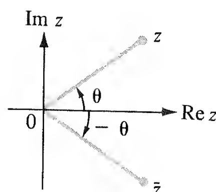


Figura A.4

$$\arg \bar{z} = -\arg z \tag{12}$$

Podemos usar $|z|$ y arg z para describir una forma que frecuentemente, es más conveniente para representar a los números complejos*. De la Figura A.3 es evidente que si $z = \alpha + i\beta$, $r = |z|$, y $\theta = \arg z$, entonces

$$\alpha = r\cos\theta \quad \text{y} \quad \beta = r \operatorname{sen}\theta \tag{13}$$

Veremos al final de este apéndice que

$$e^{i\theta} = \cos\theta + i \operatorname{sen}\theta \tag{14}$$

Como $\cos(-\theta) = \cos\theta$ y $\operatorname{sen}(-\theta) = -\operatorname{sen}\theta$, tenemos también

$$e^{-i\theta} = \cos(-\theta) + i\operatorname{sen}(-\theta) = \cos\theta - i\operatorname{sen}\theta \tag{14'}$$

La Fórmula (14) se conoce como la *fórmula de Euler*†. Usando esta fórmula y la Ecuación (13), tenemos

$$z = \alpha + i\beta = r\cos\theta + ir\operatorname{sen}\theta = r(\cos\theta + i\operatorname{sen}\theta)$$

o bien

$$z = re^{i\theta} \tag{15}$$

La representación (15) se conoce como la *forma polar* del número complejo z.

Ejemplo 5 Determinar la forma polar de los siguientes números complejos (i) 1, (ii) −1, (iii) i, (iv) $1 + i$, (v) $-1 - \sqrt{3}i$ y (vi) $-2 + 7i$.

Solución Los seis puntos están representados en la Figura A.5.

i. De la Figura A.5a, es claro que arg 1 = 0. Como Re 1 = 1, vemos que, en la forma polar, $1 = 1/e^{i0} = 1/e^0 = 1$.

* Esta representación es muy familiar para aquéllos que hayan estudiado coordenadas polares en un curso de cálculo.

† Llamada así en honor del gran matemático suizo Leonhard Euler (1707-1783).

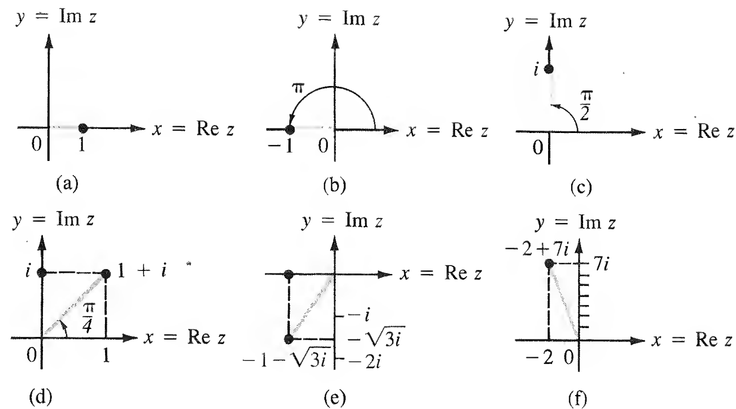


Figura A.5

ii. Ya que $\arg(-1) = \pi$ (Figura A.5b) y $|-1| = 1$, se tiene que

$$-1 = 1e^{\pi i} = e^{i\pi}$$

iii. De la Figura A5c se ve que $\arg i = \pi/2$. Como $|i| = \sqrt{0^2 + 1^2} = 1$, se deduce que

$$i = e^{i\pi/2}$$

iv. $\arg(1 + i) = \tan^{-1}(1/1) = (\pi/4)$ y $|1 + i| = \sqrt{1^2 + 1^2} = \sqrt{2}$, de modo que

$$1 + i = \sqrt{2}e^{i\pi/4}$$

v. Aquí $\tan^{-1}(\beta/\alpha) = \tan^{-1}\sqrt{3} = \pi/3$. Sin embargo, $\arg z$ está en el tercer cuadrante, así que $\theta = \pi/3 + \pi = 4\pi/3$. Tenemos además que $|-1 - \sqrt{3}| = \sqrt{1^2 + (\sqrt{3})^2} = \sqrt{1 + 3} = 2$, de modo que

$$-1 - \sqrt{3} = 2e^{4\pi i/3}$$

vi. Para resolver este problema necesitamos una calculadora. Se halla que

$$\arg z = \tan^{-1}(-\tfrac{7}{2}) = \tan^{-1}(-3.5) \approx -1.2925.$$

Pero $\tan^{-1} x$ está definido como número en el intervalo $(-\pi/2, \pi/2)$. De la Figura A.5f, θ está en el segundo cuadrante, y por lo tanto, $\arg z = \tan^{-1}(-3.5) + \pi \approx 1.8491$. Por otra parte,

$$|-2 + 7i| = \sqrt{(-2)^2 + 7^2} = \sqrt{53}.$$

En consecuencia,

$$-2 + 7i \approx \sqrt{53}e^{1.8491i}$$

Ejemplo 6 Convierta los siguientes números complejos de forma polar a forma cartesiana: (i) $2e^{i\pi/3}$; (ii) $4e^{3\pi i/2}$.

Solución **i.** $e^{i\pi/3} = \cos\pi/3 + i\,\text{sen}\,\pi/3 = \frac{1}{2} + (\sqrt{3}/2)i$. Así $2e^{i\pi/3} = 1 + \sqrt{3}i$.
ii. $e^{3\pi i/2} = \cos 3\pi/2 + i\,\text{sen}\,3\pi/2 = 0 + i(-1) = -i$. Así $4e^{3\pi i/2} = -4i$.

Si $\theta = \arg z$, entonces, por la Ecuación (12), $\arg \bar{z} = -\theta$. Así, como $|\bar{z}| = |z|$:

Si $z = re^{i\theta}$, entonces $\bar{z} = re^{-i\theta}$	(16)

Suponiendo que se expresa un número complejo en su forma polar $z = re^{i\theta}$. Entonces

$$z^n = (re^{i\theta})^n = r^n(e^{i\theta})^n = r^n e^{in\theta} = r^n(\cos n\theta + i\,\text{sen}\,n\theta) \quad (17)$$

La fórmula (17) es útil en una gran variedad de cálculos. En particular, cuando $r = |z| = 1$, se obtiene la *fórmula de De Moivre.**

$(\cos\theta + i\,\text{sen}\,\theta)^n = \cos n\theta + i\,\text{sen}\,n\theta$	(18)

Ejemplo 7 Calcule $(1 + i)^5$.

Solución En el Ejemplo 5(iv) se demostró que $1 + i = \sqrt{2}e^{\pi i/4}$. Entonces

$$(1 + i)^5 = (\sqrt{2}e^{\pi i/4})^5 = (\sqrt{2})^5 e^{5\pi i/4} = 4\sqrt{2}\left(\cos\frac{5\pi}{4} + i\,\text{sen}\frac{5\pi}{4}\right)$$

$$= 4\sqrt{2}\left(-\frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2}}i\right) = -4 - 4i$$

Esto puede ser verificado por cálculo directo. Si tal operación no fuese tanto más difícil, trátese de evaluar $(1 + i)^{20}$ en forma directa. Procediendo como antes, obtenemos

$$(1+i)^{20} = (\sqrt{2})^{20}e^{20\pi i/4} = 2^{10}(\cos 5\pi + i\,\text{sen}\,5\pi)$$
$$= 2^{10}(-1+0) = -1024$$

Demostración de la Fórmula de Euler‡ Probaremos que

$$e^{i\theta} = \cos\theta + i\,\text{sen}\,\theta \quad (19)$$

usando series de potencias. Si esto no le es bien conocido omita la demostración. Tenemos que

---

* Abraham De Moivre (1667-1754) fue un matemático francés muy conocido por su trabajo en teoría de la probabilidad, en series infinitas y en trigonometría. Su reputación era tan elevada que frecuentemente Newton decía a quienes recurrían a él en cuestiones de matemáticas: "Vaya usted con De Moivre; él sabe de estas cosas mejor que yo".

‡ Si el tiempo lo permite.

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots \dagger \tag{20}$$

$$\operatorname{sen} x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots \tag{21}$$

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots \tag{22}$$

Entonces
$$e^{i\theta} = 1 + (i\theta) + \frac{(i\theta)^2}{2!} + \frac{(i\theta)^3}{3!} + \frac{(i\theta)^4}{4!} + \frac{(i\theta)^5}{5!} + \cdots \tag{23}$$

Ahora $i^2 = -1$, $i^3 = -i$, $i^4 = 1$, $i^5 = i$, y así sucesivamente. Así, (23) puede escribirse

$$\begin{aligned} e^{i\theta} &= 1 + i\theta - \frac{\theta^2}{2!} - \frac{i\theta^3}{3!} + \frac{\theta^4}{4!} + \frac{i\theta^5}{5!} - \cdots \\ &= \left(1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \cdots\right) + i\left(\theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \cdots\right) \\ &= \cos\theta + i \operatorname{sen}\theta \end{aligned}$$

Esto completa la demostración. ■

Problemas A.2

En los Problemas del 1 al 5 efectúe la operación indicada.

1. $(2-3i)+(7-4i)$
2. $3(4+i)-5(-3+6i)$
3. $(1+i)(1-i)$
4. $(2-3i)(4+7i)$
5. $(-3+2i)(7+3i)$

En los Problemas del 6 al 15 convierta el número complejo a su forma polar.

6. $5i$
7. $5+5i$
8. $-2-2i$
9. $3-3i$
10. $2+2\sqrt{3}i$
11. $3\sqrt{3}+3i$
12. $1-\sqrt{3}i$
13. $4\sqrt{3}-4i$
14. $-6\sqrt{3}-6i$
15. $-1-\sqrt{3}i$

En los Problemas del 16 al 25 convierta de la forma polar a la forma cartesiana.

16. $e^{3\pi i}$
17. $2e^{-7\pi i}$
18. $\frac{1}{2}e^{3\pi i/4}$
19. $\frac{1}{2}e^{-3\pi i/4}$
20. $6e^{\pi i/6}$
21. $4e^{5\pi i/6}$
22. $4e^{-5\pi i/6}$
23. $3e^{-2\pi i/3}$
24. $\sqrt{3}e^{23\pi i/4}$
25. e^{i}

En los Problemas del 26 al 34 calcule el conjugado del número dado.

26. $3-4i$
27. $4+6i$
28. $-3+8i$
29. $-7i$
30. 16
31. $2e^{\pi i/7}$
32. $4e^{3\pi i/5}$
33. $3e^{-4\pi i/11}$
34. $e^{0.012i}$

† Aunque no lo probamos aquí, esta expansión en series es válida también cuando x es un número complejo.

35. Muestre que $z = \alpha + i\beta$ es real si y sólo si $z = \bar{z}$. [*Sugerencia:* Si $z = \bar{z}$ muestre que $\beta = 0$.]

36. Muestre que $z = \alpha + i\beta$ es imaginario puro si y sólo si $z = -\bar{z}$. [*Sugerencia:* Si $z = -\bar{z}$, muestre que $\alpha = 0$.]

37. Para cualquier número complejo z muestre que $z\bar{z} = |z|^2$.

38. Muestre que el círculo de radio 1 centrado en el origen (el *círculo unitario*) es el conjunto de puntos en el plano complejo que satisface $|z| = 1$.

39. Para cualesquiera números z_0 complejo y a real describa $\{z: |z - z_0| = a\}$.

40. Describa $\{z: |z - z_0| \le a\}$, donde z_0 y a son como en el Problema 39.

★ **41.** Sea $p(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + a_{n-2}\lambda^{n-2} + \cdots + a_1\lambda + a_0$ con $a_0, a_1, \ldots, a_{n-1}$ números reales. Muestre que si $p(z) = 0$, entonces $p(\bar{z}) = 0$. Esto es: *Las raíces de un polinomio con coeficientes reales ocurren en pares de conjugados complejos.*

42. Derive expresiones para $\cos 4\theta$ y $\operatorname{sen} 4\theta$ comparando la fórmula de De Moivre y la expansión de $(\cos\theta + i \operatorname{sen}\theta)^4$.

43. Demuestre la fórmula de De Moivre por inducción matemática. [*Sugerencia:* Recuerde las identidades trigonométricas $\cos(x+y) = \cos x \cos y - \operatorname{sen} x \operatorname{sen} y$ y $\operatorname{sen}(x+y) = \operatorname{sen} x \cos y + \cos x \operatorname{sen} y$.]

Respuestas a los problemas de número impar

Capítulo 1

Problemas 1.2

1. $x_1 = \frac{-13}{5}$, $x_2 = \frac{-11}{5}$; det $= -10$

3. no tiene solución; det $= 0$

5. $x_1 = \frac{11}{2}$, $x_2 = -30$; det $= -2$

7. número infinito de soluciones; $x_2 = \frac{2}{3}x_1$, en donde x_1 es arbitraria; det $= 0$

9. $x_1 = -1$, $x_2 = 2$; det $= -1$

11. det $= a^2 - b^2$; si $a^2 - b^2 \neq 0$ (es decir, si $a \neq \pm b$), entonces $x_1 = x_2 = c/(a+b)$. Si $a^2 - b^2 = 0$, entonces $a = \pm b$. Si $a \neq 0$ y $a = b$, entonces hay un número infinito de soluciones dadas por $x_2 = c/a - x_1$. Si $a \neq 0$ y $a = -b$, entonces no hay soluciones.

13. det $= -2ab$; así que hay una solución única si a y b son distintos de cero.

15. $a = b = 0$ y $c \neq 0$ o bien $d \neq 0$

17. no hay punto de intersección

19. Las rectas coinciden. Cualquier punto de la forma $(x, (4x - 10)/6)$ es un punto de intersección.

21. $(\frac{67}{45}, \frac{2}{15})$ **23.** $\sqrt{13}/13$

25. $\sqrt{61}/5$ **27.** $\sqrt{5}$

29. Como la pendiente de la recta dada L es $-\frac{a}{b}$, la pendiente de $L_\perp$ es $\frac{b}{a}$. La ecuación de la recta $L_\perp$ perpendicular a L y que pasa por (x_1, y_1) está dada por $\frac{y - y_1}{x - x_1} = \frac{b}{a}$, o bien por $bx - ay = bx_1 - ay_1$. El único punto de intersección de L y $L_\perp$ es

$$(x_0, y_0) = \left(\frac{bc - abx_1 + a^2y_1}{a^2 + b^2}, \frac{ac - aby_1 + b^2x_1}{a^2 + b^2}\right).$$

Entonces d es la distancia entre (x_0, y_0) y (x_1, y_1) y después de algunas operaciones

$$d^2 = \frac{1}{(a^2 + b^2)^2} \times (a^2c^2 - 2a^2bcy_1 + a^2b^2y_1^2 - 2a^3cx_1 + 2a^3bx_1y_1 + a^4x_1^2 + c^2b^2 - 2ab^2cx_1 + a^2b^2x_1^2 - 2b^3cy_1 + 2ab^3x_1y_1 + b^4y_1^2)$$

$$= \frac{a^2 + b^2}{(a^2 + b^2)^2}(c^2 - 2abcy_1 + b^2y_1^2 - 2acx_1 + 2abx_1y_1 + a^2x_1^2) = \frac{1}{(a^2 + b^2)}(ax_1 + by_1 - c)^2.$$

Por lo tanto

$$d = \frac{|ax_1 + by_1 - c|}{\sqrt{a^2 + b^2}}.$$

35. Número infinito de soluciones; x = número de tazas; y = número de platos; las soluciones son $(x, 240 - \frac{3}{2}x)$.

37. 32 refrescos; 128 malteadas.

Problemas 1.3

1. $\begin{pmatrix} 2 \\ -3 \\ 11 \end{pmatrix}$ **3.** $\begin{pmatrix} -4 \\ 0 \\ 4 \end{pmatrix}$

5. $\begin{pmatrix} -31 \\ 22 \\ -27 \end{pmatrix}$ **7.** $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$

9. $\begin{pmatrix} -11 \\ 11 \\ -10 \end{pmatrix}$ **11.** $(1, 2, 5, 7)$

13. $(-8, 12, 4, 20)$

15. $(8, -5, 7, -1)$

17. $(7, 2, 4, 11)$

19. $(-11, 9, 18, 18)$

21. $\mathbf{a}+\mathbf{0}=\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix}+\begin{pmatrix}0\\0\\\vdots\\0\end{pmatrix}$

$=\begin{pmatrix}a_1+0\\a_2+0\\\vdots\\a_n+0\end{pmatrix}=\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix}=\mathbf{a}$

23. $\mathbf{a}+\mathbf{b}=\begin{pmatrix}a_1+b_1\\a_2+b_2\\\vdots\\a_n+b_n\end{pmatrix},$

$\alpha(\mathbf{a}+\mathbf{b})=\begin{pmatrix}\alpha(a_1+b_1)\\\alpha(a_2+b_2)\\\vdots\\\alpha(a_n+b_n)\end{pmatrix}$

$=\begin{pmatrix}\alpha a_1+\alpha b_1\\\alpha a_2+\alpha b_2\\\vdots\\\alpha a_n+\alpha b_n\end{pmatrix}$

$=\begin{pmatrix}\alpha a_1\\\alpha a_2\\\vdots\\\alpha a_n\end{pmatrix}+\begin{pmatrix}\alpha b_1\\\alpha b_2\\\vdots\\\alpha b_n\end{pmatrix}$

$=\alpha\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix}+\alpha\begin{pmatrix}b_1\\b_2\\\vdots\\b_n\end{pmatrix}$

$=\alpha\mathbf{a}+\alpha\mathbf{b};$

$(\alpha+\beta)\mathbf{a}=\begin{pmatrix}(\alpha+\beta)a_1\\(\alpha+\beta)a_2\\\vdots\\(\alpha+\beta)a_n\end{pmatrix}$

$=\begin{pmatrix}\alpha a_1+\beta a_1\\\alpha a_2+\beta a_2\\\vdots\\\alpha a_n+\beta a_n\end{pmatrix}$

$=\begin{pmatrix}\alpha a_1\\\alpha a_2\\\vdots\\\alpha a_n\end{pmatrix}+\begin{pmatrix}\beta a_1\\\beta a_2\\\vdots\\\beta a_n\end{pmatrix}$

$=\alpha\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix}+\beta\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix}$

$=\alpha\mathbf{a}+\beta\mathbf{a};$

$(\alpha\beta)\mathbf{a}=\begin{pmatrix}\alpha\beta a_1\\\alpha\beta a_2\\\vdots\\\alpha\beta a_n\end{pmatrix}=\alpha\begin{pmatrix}\beta a_1\\\beta a_2\\\vdots\\\beta a_n\end{pmatrix}$

$=\alpha\left[\beta\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix}\right]=\alpha(\beta\mathbf{a})$

25. $\mathbf{d}_1+\mathbf{d}_2$ representa la demanda combinada de las dos fábricas para cada una de las cuatro materias primas que se necesitan para producir una unidad de cada producto; $2\mathbf{d}_1$ representa la demanda de la fábrica 1 para cada una de las cuatro materias primas que se necesitan para producir dos unidades de su producto.

27. $\mathbf{w}=\begin{pmatrix}3\\0\\5\end{pmatrix}$

Problemas 1.4

1. $\begin{pmatrix}3&9\\6&15\\-3&6\end{pmatrix}$ **3.** $\begin{pmatrix}2&2\\-2&-1\\6&-1\end{pmatrix}$

5. $\begin{pmatrix}0&0\\0&0\\0&0\end{pmatrix}$ **7.** $\begin{pmatrix}-2&4\\7&15\\-15&10\end{pmatrix}$

9. $\begin{pmatrix}4&10\\17&22\\-9&1\end{pmatrix}$ **11.** $\begin{pmatrix}0&6\\5&14\\-9&9\end{pmatrix}$

13. $\begin{pmatrix}1&-5&0\\-3&4&-5\\-14&13&-1\end{pmatrix}$

15. $\begin{pmatrix}1&1&5\\9&5&10\\7&-7&3\end{pmatrix}$

17. $\begin{pmatrix}-1&-1&-1\\-3&-3&-10\\-7&3&5\end{pmatrix}$

19. $\begin{pmatrix}-1&-1&-5\\-9&-5&-10\\-7&7&-3\end{pmatrix}$

25. $\begin{pmatrix}1&1&1&0\\1&1&1&0\\1&1&1&1\\0&0&1&1\end{pmatrix}$

Problemas 1.5

1. -14 **3.** 1 **5.** $ac+bd$

7. 51 **9.** $a=0$ **11.** 4

13. 28

15. $\begin{pmatrix}8&20\\-4&11\end{pmatrix}$ **17.** $\begin{pmatrix}-3&-3\\1&3\end{pmatrix}$

19. $\begin{pmatrix}13&35&18\\20&26&20\end{pmatrix}$

21. $\begin{pmatrix}19&-17&34\\8&-12&20\\-8&-11&7\end{pmatrix}$

23. $\begin{pmatrix}18&15&35\\9&21&13\\10&9&9\end{pmatrix}$ **25.** $(7\ \ 16)$

27. $\begin{pmatrix}3&-2&1\\4&0&6\\5&1&9\end{pmatrix}$ **29.** $\begin{pmatrix}a&b&c\\d&e&f\\g&h&j\end{pmatrix}$

31. Si $D=a_{11}a_{22}-a_{12}a_{21}$, entonces $\begin{pmatrix}b_{11}&b_{21}\\b_{21}&b_{22}\end{pmatrix}$

$=\begin{pmatrix}a_{22}/D&-a_{12}/D\\-a_{21}/D&a_{11}/D\end{pmatrix}$

33. (a) 3 en el grupo 1, 4 en el grupo 2, 5 en el grupo 3

(b) $\begin{pmatrix}2&1&1&0&0\\1&1&0&1&0\\1&0&2&0&1\end{pmatrix}$

35. son ortogonales

37. son ortogonales

39. son ortogonales

41. todas las α y β que satisfacen $5\alpha+4\beta=25$ ($\beta=(25-5\alpha)/4$, α es arbitraria)

43. (a) (2, 3, 5, 1)

(b) $\begin{pmatrix}1\\\frac{3}{2}\\\frac{1}{2}\\2\end{pmatrix}$ (c) 11

45. (a) $\begin{pmatrix}80{,}000&45{,}000&40{,}000\\50&20&10\end{pmatrix}$

(b) $\begin{pmatrix}1\\3\\1\end{pmatrix}$ (c) Dinero: 255,000; Acciones: 120

47. $\begin{pmatrix}0&-8\\32&32\end{pmatrix}$ **49.** $\begin{pmatrix}11&38\\57&106\end{pmatrix}$

51. $A^2=\begin{pmatrix}0&0&1&0&0\\0&0&0&1&0\\0&0&0&0&1\\0&0&0&0&0\\0&0&0&0&0\end{pmatrix},$

$A^3=\begin{pmatrix}0&0&0&1&0\\0&0&0&0&1\\0&0&0&0&0\\0&0&0&0&0\\0&0&0&0&0\end{pmatrix},$

$A^4=\begin{pmatrix}0&0&0&0&1\\0&0&0&0&0\\0&0&0&0&0\\0&0&0&0&0\\0&0&0&0&0\end{pmatrix},$

$A^5=\begin{pmatrix}0&0&0&0&0\\0&0&0&0&0\\0&0&0&0&0\\0&0&0&0&0\\0&0&0&0&0\end{pmatrix}$

53. $PQ=\begin{pmatrix}\frac{11}{90}&\frac{41}{90}&\frac{19}{45}\\\frac{11}{120}&\frac{71}{120}&\frac{19}{60}\\\frac{1}{5}&\frac{1}{5}&\frac{3}{5}\end{pmatrix};$

todos los elementos son no negativos y $\frac{11}{90}+\frac{41}{90}+\frac{19}{45}=\frac{11}{120}+\frac{71}{120}+\frac{19}{60}=\frac{1}{5}+\frac{1}{5}+\frac{3}{5}=1.$

55. Sean $P=(p_{ij})$ y $Q=(q_{ij})$ matrices de probabilidad de $k\times k$. Sea $PQ=C=(c_{ij})$. La suma de los elementos en el renglón m de PQ es

$c_{m1}+c_{m2}+c_{m3}+\cdots$
$+c_{mk}=p_{m1}q_{11}+p_{m2}q_{21}$
$+p_{m3}q_{31}+\cdots+p_{mk}q_{k1}$
$+p_{m1}q_{12}+p_{m2}q_{22}+p_{m3}q_{32}$
$+\cdots+p_{mk}q_{k2}+p_{m1}q_{13}$
$+p_{m2}q_{23}+p_{m3}p_{33}+\cdots$
$+p_{mk}q_{k3}$
$\vdots$
$+p_{m1}q_{1k}+p_{m2}q_{2k}+p_{m3}q_{3k}$
$+\cdots+p_{mk}q_{kk}$

(los elementos entre paréntesis son los de un renglón de Q, cuya suma es 1)

$\downarrow$
$=p_{m1}(q_{11}+q_{12}+q_{13}+\cdots$
$+q_{1k})+p_{m2}(q_{21}+q_{22}$
$+q_{23}+\cdots+q_{2k})$
$+p_{m3}(q_{31}+q_{32}+q_{33}+\cdots$
$+q_{3k})+\cdots+p_{mk}(q_{k1}+q_{k2}$
$+q_{k3}+\cdots+q_{kk})$
$=p_{m1}(1)+p_{m2}(1)+p_{m3}(1)$
$+\cdots+p_{mk}(1)=1.$

57. (a) jugador 2 > jugador 4 > jugador 1 > jugador 3

(b) puntuación = número de juegos que ganó más la mitad del número de juegos ganados por cada jugador que el jugador dado haya derrotado.

59. $A(B+C)=\begin{pmatrix}1&2&4\\3&-1&0\end{pmatrix}\times\begin{pmatrix}1&9\\2&11\\10&1\end{pmatrix}$

$=\begin{pmatrix}45&35\\1&16\end{pmatrix};$

$AB+AC=\begin{pmatrix}24&14\\7&17\end{pmatrix}+\begin{pmatrix}21&20\\-6&-1\end{pmatrix}$

$=\begin{pmatrix}45&35\\1&16\end{pmatrix}$

61. 36 **63.** 9840

65. $\frac{13}{3}+\frac{15}{4}+\frac{17}{5}=\frac{689}{60}$

67. $(1^2+2^2+3^2)(2^3+3^3+4^3)=1386$

69. $\sum_{k=0}^{5}(-3)^k$ **71.** $\sum_{k=1}^{n}k^{1/k}$

73. $\sum_{k=0}^{9}\frac{(-1)^{k+1}}{a^k}$

75. $\sum_{k=2}^{7}k^2\cdot 2k=\sum_{k=2}^{7}2k^3$

77. $\sum_{i=1}^{3}\sum_{j=1}^{2}a_{ij}$ **79.** $\sum_{k=1}^{5}a_{3k}b_{k2}$

Problemas 1.6

Nota: Donde haya habido un número infinito de soluciones, se presentan las soluciones con la última variable elegida arbitrariamente. Las soluciones pueden presentarse también de otras maneras.

1. $(2,-3,1)$

3. $(3+\frac{2}{9}x_3,\frac{8}{9}x_3,x_3)$, x_3 arbitrario.

5. $(-9,30,14)$

7. no hay solución.

9. $(-\frac{4}{5}x_3,\frac{9}{5}x_3,x_3)$, x_3 es arbitraria

11. $(-1,\frac{5}{2}+\frac{1}{2}x_3,x_3)$, x_3 es arbitraria

13. no hay solución

15. $(\frac{20}{13}-\frac{4}{13}x_4,\frac{-28}{13}+\frac{3}{13}x_4,\frac{-45}{13}+\frac{9}{13}x_4,x_4)$, es arbitraria

17. $(18-4x_4,\frac{-15}{2}+2x_4,-31+7x_4,x_4)$, x_4 es arbitraria

19. no hay solución

21. forma escalonada (por renglones)

23. forma escalonada reducida

25. de ninguna de esas clases

27. forma escalonada reducida

29. de ninguna de esas clases

31. forma escalonada: $\begin{pmatrix}1&-6\\0&1\end{pmatrix}$; forma escalonada reducida: $\begin{pmatrix}1&0\\0&1\end{pmatrix}$

33. forma escalonada: $\begin{pmatrix}1&-2&4\\0&1&-\frac{4}{11}\\0&0&1\end{pmatrix};$

forma escalonada reducida:
$$\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}$$

35. forma escalonada:
$$\begin{pmatrix}1&-\frac{7}{2}\\0&1\\0&0\end{pmatrix};$$
forma escalonada reducida
$$\begin{pmatrix}1&0\\0&1\\0&0\end{pmatrix}$$

37. $x_1 = 30{,}000 - 5x_3$
$x_2 = x_3 - 5000$
$5000 \le x_3 \le 6000$; no

39. No hay solución única (2 ecuaciones en 3 incógnitas); si tiene 200 acciones de McDonald's, entonces 100 acciones de Hilton y 300 acciones de Eastern.

41. La forma escalonada de la matriz aumentada que representa este sistema es
$$\left(\begin{array}{ccc|c}1&-\frac{1}{2}&\frac{3}{2}&a/2\\0&1&\frac{-19}{5}&\frac{2}{5}(b-\frac{3}{2}a)\\0&0&0&-2a+3b+c\end{array}\right)$$
el cual es inconsistente si $-2a+3b+c\neq 0$ o bien $c\neq 2a-3b$.

43. $a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{32}a_{21}-a_{13}a_{22}a_{31}-a_{12}a_{21}a_{33}-a_{11}a_{32}a_{23}\neq 0$

45. (1.900812947, 4.194110816, − 11.34851834)

47. $\begin{pmatrix}2&-1\\4&5\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}=\begin{pmatrix}3\\7\end{pmatrix}$

49. $\begin{pmatrix}3&6&-7\\2&-1&3\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}0\\1\end{pmatrix}$

51. $\begin{pmatrix}0&1&-1\\1&0&1\\3&2&0\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}7\\2\\-5\end{pmatrix}$

Problemas 1.7

1. (0, 0) **3.** (0, 0, 0)

5. $(\frac{1}{6}x_3, \frac{5}{6}x_3, x_3)$, x_3 es arbitraria

7. (0, 0)

9. $(-4x_4, 2x_4, 7x_4, x_4)$, x_4 es arbitraria

11. (0, 0) **13.** (0, 0, 0)

15. $k = \frac{95}{11}$

17. La solución más simple a la ecuación no homogénea se obtiene aplicando $x_2 = 0$. Entonces la solución general es $(2, 0) + x_2(3, 1)$; x_2 es arbitraria.

19. Si $x_3 = 0$, una solución no homogénea es (2, 0, 0) y la solución general es $(2, 0, 0) + x_3(-\frac{1}{3}, -\frac{4}{3}, 1)$; x_3 es arbitraria.

21. Si $x_3 = x_4 = 0$, una solución no homogénea es (−1, 4, 0,0) y la solución general es $(-1, 4, 0, 0) + x_3(-3, 4, 1, 0) + x_4(5, -7, 0, 1)$

23. $(c_1y_1 + c_2y_2)'' + a(x)(c_1y_1 + c_2y_2)' + b(x)(c_1y_1 + c_2y_2)$
$= c_1y_1'' + c_2y_2'' + a(x)c_1y_1' + a(x)c_2y_2' + b(x)c_1y_1 + b(x)c_2y_2$
$= c_1(y_1'' + a(x)y_1' + b(x)y_1) + c_2(y_2'' + a(x)y_2' + b(x)y_2)$
$= c_1\cdot 0 + c_2\cdot 0 = 0$,
puesto que y_1 y y_2 son la solución de (7)

Problemas 1.8

1. $\begin{pmatrix}2&-1\\-3&2\end{pmatrix}$ **3.** $\begin{pmatrix}0&1\\1&0\end{pmatrix}$

5. no invertible

7. $\begin{pmatrix}\frac{1}{3}&-\frac{1}{3}&-\frac{1}{3}\\0&\frac{1}{2}&1\\0&0&-1\end{pmatrix}$

9. no invertible

11. no invertible

13. $\begin{pmatrix}\frac{7}{3}&-\frac{1}{3}&-\frac{1}{3}&-\frac{2}{3}\\\frac{4}{9}&-\frac{1}{9}&-\frac{4}{9}&\frac{1}{9}\\-\frac{1}{9}&-\frac{2}{9}&\frac{1}{9}&\frac{2}{9}\\-\frac{5}{3}&\frac{2}{3}&\frac{2}{3}&\frac{1}{3}\end{pmatrix}$

15. $\begin{pmatrix}0&1&0&2\\1&-1&-2&2\\0&1&3&-3\\-2&2&3&-2\end{pmatrix}$

17. $(A_1A_2\cdots A_m)^{-1} = A_m^{-1}A_{m-1}^{-1}\cdots A_2^{-1}A_1^{-1}$ puesto que $(A_m^{-1}A_{m-1}^{-1}\cdots A_2^{-1}A_1^{-1})(A_1A_2\cdots A_{m-1}A_m) = (A_m^{-1}A_{m-1}^{-1}\cdots A_2^{-1})(A_1^{-1}A_1)\times A_2\cdots A_{m-1}A_m = (A_m^{-1}A_{m-1}^{-1}\cdots A_2^{-1})\times(A_2\cdots A_{m-1}A_m) = \cdots = I$

19. $A^{-1} = \dfrac{1}{a_{11}a_{22}-a_{21}a_{12}}\times\begin{pmatrix}a_{22}&-a_{12}\\-a_{21}&a_{11}\end{pmatrix}$. Si $A = \pm I$, entonces $A^{-1} = A$. Si $a_{11} = -a_{22}$ y $a_{21}a_{12} = 1 - a_{11}^2$, entonces $a_{11}a_{22} - a_{21}a_{12} = -a_{11}^2 - (1 - a_{11}^2) = -1$. Por tanto,
$$A^{-1} = \begin{pmatrix}-a_{22}&a_{12}\\a_{21}&-a_{11}\end{pmatrix} = \begin{pmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{pmatrix} = A.$$

21. El sistema $B\mathbf{x} = \mathbf{0}$ tiene un número infinito de soluciones (por el Teorema 1.7.1). Pero si $B\mathbf{x} = \mathbf{0}$, entonces $AB\mathbf{x} = \mathbf{0}$. Así, del Teorema 6 (partes [i] y [ii], AB es no invertible.

23. $\begin{pmatrix}\text{sen}\,\theta&\cos\theta&0\\\cos\theta&-\text{sen}\,\theta&0\\0&0&1\end{pmatrix}$

es su propia inversa (ya que $\text{sen}^2\theta + \cos^2\theta = 1$).

25. Si la i-ésima componente de la diagonal es cero, entonces en la reducción de renglón de A el renglón i es cero, de manera que, por lo establecido en el Paso 3(b) de la página 59, A no es invertible. Por otra parte, si
$$A = \text{diag}(a_1, a_2, \ldots, a_n)$$
entonces
$$A^{-1} = \text{diag}\left(\frac{1}{a_1}, \frac{1}{a_2}, \ldots, \frac{1}{a_n}\right).$$

27. $\begin{pmatrix}\frac{1}{2}&-\frac{1}{6}&\frac{7}{30}\\0&\frac{1}{3}&-\frac{4}{15}\\0&0&\frac{1}{5}\end{pmatrix}$

29. Se demostró que el resultado en el caso A es triangular superior. La demostración en el caso triangular inferior es semejante. Considere el sistema homogéneo.
$$\begin{pmatrix}a_{11}&a_{12}&a_{13}&\cdots&a_{1,n-1}&a_{1n}\\0&a_{22}&a_{23}&\cdots&a_{2,n-1}&a_{2n}\\\vdots&\vdots&\vdots&&\vdots&\vdots\\0&0&0&\cdots&a_{n-1,n-1}&a_{n-1,n}\\0&0&0&\cdots&0&a_{nn}\end{pmatrix}\times\begin{pmatrix}x_1\\x_2\\\vdots\\x_{n-1}\\x_n\end{pmatrix}=\begin{pmatrix}0\\0\\\vdots\\0\\0\end{pmatrix}$$
Supóngase que $a_{11}, a_{22}, \ldots, a_{nn}$ son todos diferentes de cero. La última ecuación en el sistema homogéneo es $a_{nn}x_n = 0$ y, como $a_{nn}\neq 0$, entonces $x_n = 0$. La penúltima ecuación es
$$a_{n-1,n-1}x_{n-1} + a_{n-1,n}x_n = 0$$
y $a_{n-1,n-1}\neq 0$, $x_n = 0$ implica que $x_{n-1} = 0$. De manera similar, concluimos que $x_1 = x_2 = \cdots = x_{n-1} = x_n = 0$, de modo que la única solución al sistema homogéneo es la solución trivial. Por el Teorema 6 (partes (i) y (ii), A es invertible. Recíprocamente, supóngase una de las componentes de la diagonal, por ejemplo a_{11}, es igual a cero. Entonces el sistema homogéneo $A\mathbf{x} = \mathbf{0}$ tiene la solución
$$\mathbf{x} = \begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}$$
[Si $a_{jj} = 0$ con $j\neq 1$, entonces se elige $\mathbf{x}$ como el vector con 1 en la posición j y cero en cualquier otro lugar.] Usando nuevamente el Teorema 6, concluimos que A no es invertible.

31. cualquier múltiplo no cero de (1, 2)

33. 3 sillas y 2 mesas

35. 4 unidades de A y 5 unidades de B

37. (a) $A = \begin{pmatrix}0.293&0&0\\0.014&0.207&0.017\\0.044&0.010&0.216\end{pmatrix}$;

$I - A$
$$= \begin{pmatrix}0.707&0&0\\-0.014&0.793&-0.017\\-0.044&-0.010&0.784\end{pmatrix}$$

(b) $\begin{pmatrix}195492.2207\\25932.85859\\13580.33966\end{pmatrix}$

39. $\begin{pmatrix}1&\frac{1}{2}\\0&1\end{pmatrix}$; sí

41. $\begin{pmatrix}1&\frac{2}{3}&\frac{1}{3}\\0&1&1\\0&0&1\end{pmatrix}$; sí

43. $\begin{pmatrix}1&-\frac{1}{2}&2\\0&1&-14\\0&0&0\end{pmatrix}$; no

45. $\begin{pmatrix}1&0&2&3\\0&1&2&7\\0&0&1&\frac{10}{7}\\0&0&0&0\end{pmatrix}$; no

Problemas 1.9

1. $\begin{pmatrix}-1&6\\4&5\end{pmatrix}$ **3.** $\begin{pmatrix}2&-1&1\\3&2&4\end{pmatrix}$

5. $\begin{pmatrix}1&-1&1\\2&0&5\\3&4&5\end{pmatrix}$ **7.** $\begin{pmatrix}1&0\\0&1\\1&0\\0&1\end{pmatrix}$

9. $\begin{pmatrix}a&d&g\\b&e&h\\c&f&j\end{pmatrix}$

11. $[(A + B)^t]_{ij} = (A + B)_{ji} = a_{ji} + b_{ji} = (A^t)_{ij} + (B^t)_{ij}$. Por consiguiente, la componente ij de $(A + B)^t$ es igual a la componente ij de A^t más la componente ij de B^t.

15. Si A es de $m\times n$, entonces A^t es de $n\times m$ y AA^t es de $m\times m$. También, $(AA^t)^t = (A^t)^tA^t = AA^t$.

17. Si A es triangular superior y $B = A^t$, entonces $b_{ij} = a_{ji} = 0$ si $j > i$. En consecuencia, B es triangular inferior.

19. $(A + B)^t = A^t + B^t = -A - B = -(A + B)$

21. $(AB)^t = B^tA^t = (-B)(-A) = (-1)^2BA = BA$

Problemas 1.10

1. Si, P_{12}.

3. No [se usan dos operaciones: P_{12} seguida de $A_{1,2}(1)$]

5. No [se usan dos operaciones: $M_1(3)$ y $M_2(3)$]

7. No [se utilizan dos operaciones: P_{13} seguida de P_{12}]

9. Si, $A_{12}(2)$

11. No [se utilizan dos operaciones: $A_{1,2}(1)$ y $A_{3,4}(1)$]

13. $\begin{pmatrix}1&0&0\\0&4&0\\0&0&1\end{pmatrix}$

15. $\begin{pmatrix}1&-3&0\\0&1&0\\0&0&1\end{pmatrix}$

17. $\begin{pmatrix}0&0&1\\0&1&0\\1&0&0\end{pmatrix}$ **19.** $\begin{pmatrix}1&0&0\\0&1&1\\0&0&1\end{pmatrix}$

21. $\begin{pmatrix}0&0\\0&-2\end{pmatrix}$ **23.** $\begin{pmatrix}1&2\\0&1\end{pmatrix}$

25. $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$ **27.** $\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

29. $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -5 & 0 & 1 \end{pmatrix}$

31. $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ **33.** $\begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{4} \end{pmatrix}$

35. $\begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

37. $\begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

39. $\begin{pmatrix} 1 & 0 & 0 & -5 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$

41. $\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 3 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}\begin{pmatrix} 1 & \frac{1}{2} \\ 0 & 1 \end{pmatrix}$

43. $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 5 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -4 \end{pmatrix}\begin{pmatrix} 1 & 0 & -\frac{1}{2} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & \frac{3}{2} \\ 0 & 0 & 1 \end{pmatrix}$

45. $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$

47. $\begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -4 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 5 \end{pmatrix}$

49. $\begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & c \end{pmatrix}\begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}$; las primeras dos matrices son elementales porque $a \neq 0$ y $c \neq 0$

51. Los casos 2×2 y 3×3 son resultado de los Problemas 49 y 50. En la respuesta al Problema 1.8.29 se demuestra el resultado. Otra comprobación puede obtenerse demostrando, como en los Problemas 49 y 50, que A puede ser escrita como el producto de matrices elementales. El paso clave es reducir A a I observando que la única vez que se divide es entre los números de la diagonal, que son diferentes de cero por hipótesis.

53. A^t es triangular superior, así que $(A^t)^{-1}$ es triangular superior por el resultado del Problema 52. Pero $(A^t)^{-1} = (A^{-1})^t$, así que $(A^{-1})^t$ es triangular superior, lo cual significa que $A^{-1} = [(A^{-1})^t]^t$ es triangular inferior.

55. $\begin{pmatrix} 2 & -3 \\ -1 & 2 \end{pmatrix}$

57. $\begin{pmatrix} \frac{13}{8} & -\frac{15}{8} & \frac{5}{4} \\ -\frac{1}{2} & \frac{1}{2} & 0 \\ -\frac{1}{8} & \frac{3}{8} & -\frac{1}{4} \end{pmatrix}$

59. Sean $B = A_{ij}$ y $D = A_{ij}A$. Entonces la kr-ésima componente d_{kr} de D está dada por

$$d_{kr} = \sum_{l=1}^{n} b_{kl}a_{lr} \qquad (*)$$

Si $k \neq j$, el renglón k de B es el renglón k de la identidad, de modo que $b_{kl} = 1$ si $l = k$, y 0 en otro caso cualquiera. Así

$$d_{kr} = b_{kk}a_{kr} = a_{kr} \text{ si } k \neq j.$$

Si $k = j$, entonces

$$b_{jl} = \begin{cases} 1, & \text{si } l = j \\ c, & \text{si } l = i \\ 0, & \text{en cualquier otro caso} \end{cases}$$

y (*) será

$$a_{jr} = b_{jj}a_{jr} + b_{ji}a_{ir} = a_{jr} + ca_{jr}$$

Así que cada componente en el renglón j de $A_{ij}A$ es la suma del componente correspondiente en el renglón j de A, y c veces el componente correspondiente en el renglón i de A.

61. $\begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix}$

63. $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$

65. $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 3 & 1 \end{pmatrix}\begin{pmatrix} 1 & -3 & 3 \\ 0 & 1 & -\frac{1}{3} \\ 0 & 0 & 0 \end{pmatrix}$

Problemas 1.11

1. $\begin{pmatrix} u & v \\ u-1 & 1+v \\ 1-u & -v \end{pmatrix}$ es la inversa derecha para cualesquier números u y v. No hay inversa izquierda.

3. Ni una ni otra

5. $\begin{pmatrix} -w+1 & -w & w \\ -z & -z+1 & z \end{pmatrix}$ es la inversa izquierda para todo w y todo z. No hay inversa derecha.

7. $\begin{pmatrix} w-2 & -2w+1 & w \\ z+\frac{3}{2} & -2z-\frac{1}{2} & z \end{pmatrix}$ es la inversa izquierda para todo w y todo z. No hay inversa derecha.

9. Cualquier vector $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ con $x + 2y + 3z = 1$ es una inversa derecha. [Aquí la identidad es la matriz (1) de 1×1.] No hay inversa izquierda.

11. Si A tiene una inversa derecha, entonces el sistema $A\mathbf{x} = \mathbf{b}$ tendría una solución para cada 3-vector $\mathbf{b}$, de acuerdo con el Teorema 1. Pero el sistema tiene solución solamente si $-b_1 + 20b_2 - 11b_3 = 0$. Por lo tanto, A no tiene inversa derecha.

13. (a) $R_1 = \begin{pmatrix} 0 & \frac{1}{4} \\ \frac{1}{3} & -\frac{1}{12} \\ 0 & 0 \end{pmatrix}$;

$R_2 = \begin{pmatrix} -\frac{1}{2} & -\frac{1}{4} \\ \frac{1}{2} & \frac{1}{12} \\ 1 & 1 \end{pmatrix}$

(b) $R_1\mathbf{b} = \begin{pmatrix} 0 \\ \frac{1}{3} \\ 0 \end{pmatrix}$; $R_2\mathbf{b} = \begin{pmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ 1 \end{pmatrix}$

15. $R = IR = (LA)R = L(AR) = LI = L$

17. (b) $L\mathbf{b} = -\frac{1}{3}\begin{pmatrix} 7 \\ 19 \end{pmatrix}$

(c) $AL\mathbf{b} = \begin{pmatrix} 27 \\ -49 \\ 26 \end{pmatrix} \neq \begin{pmatrix} 3 \\ -1 \\ 2 \end{pmatrix}$

(d) El error es el que se está razonando hacia atrás. Hemos mostrado que *si* $A\mathbf{x} = \mathbf{b}$ tiene una solución, entonces la solución es igual a $L\mathbf{b}$. Sin embargo, sabemos por el Teorema 3 que si A tiene una inversa izquierda, entonces el sistema no puede ser resuelto para todo $\mathbf{b}$. Hemos demostrado realmente que $A\mathbf{x} = \mathbf{b}$ no tiene solución si $\mathbf{b} = \begin{pmatrix} 3 \\ -1 \\ 2 \end{pmatrix}$.

Ejercicios de Repaso • Capítulo 1

1. $(\frac{1}{7}, \frac{10}{7})$ **3.** no hay solución

5. $(0, 0, 0)$ **7.** $(-\frac{1}{2}, 0, \frac{5}{2})$

9. $(\frac{1}{3}x_3, \frac{7}{3}x_3, x_3)$, x_3 es arbitraria

11. no hay solución

13. $(0, 0, 0, 0)$

15. $\begin{pmatrix} -6 & 3 \\ 0 & 12 \\ 6 & 9 \end{pmatrix}$

17. $\begin{pmatrix} 16 & 2 & 3 \\ -20 & 10 & -1 \\ -36 & 8 & 16 \end{pmatrix}$

19. $\begin{pmatrix} 17 & 39 & 41 \\ 14 & 20 & 42 \end{pmatrix}$

21. $\begin{pmatrix} 9 & 10 \\ 30 & 32 \end{pmatrix}$

23. forma escalonada reducida

25. de ninguna de esas clases

27. forma escalonada: $\begin{pmatrix} 1 & 4 & -1 \\ 0 & 1 & \frac{5}{4} \end{pmatrix}$;

forma escalonada reducida: $\begin{pmatrix} 1 & 0 & -6 \\ 0 & 1 & \frac{5}{4} \end{pmatrix}$

29. $\begin{pmatrix} 1 & \frac{3}{2} \\ 0 & 1 \end{pmatrix}$; la inversa es $\begin{pmatrix} \frac{4}{11} & -\frac{3}{11} \\ \frac{1}{11} & \frac{2}{11} \end{pmatrix}$

31. $\begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & \frac{1}{3} \\ 0 & 0 & 1 \end{pmatrix}$; la inversa es $\begin{pmatrix} -\frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{5}{8} & -\frac{1}{8} & -\frac{1}{8} \\ \frac{1}{8} & -\frac{5}{8} & \frac{3}{8} \end{pmatrix}$

33. $\begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$; la inversa es $\begin{pmatrix} \frac{5}{6} & \frac{2}{3} & -2 \\ \frac{1}{3} & \frac{2}{3} & -1 \\ -\frac{1}{6} & -\frac{1}{3} & 1 \end{pmatrix}$

35. $\begin{pmatrix} 1 & 2 & 0 \\ 2 & 1 & -1 \\ 3 & 1 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 3 \\ -1 \\ 7 \end{pmatrix}$; A^{-1} está dada en el Ejercicio 31; $x_1 = \frac{3}{4}$, $x_2 = \frac{9}{8}$, $x_3 = \frac{29}{8}$

37. $\begin{pmatrix} 2 & -1 \\ 3 & 0 \\ 1 & 2 \end{pmatrix}$; de ninguna de esas clases

39. $\begin{pmatrix} 2 & 3 & 1 \\ 3 & -6 & -5 \\ 1 & -5 & 9 \end{pmatrix}$; simétrica

41. $\begin{pmatrix} 1 & -1 & 4 & 6 \\ -1 & 2 & 5 & 7 \\ 4 & 5 & 3 & -8 \\ 6 & 7 & -8 & 9 \end{pmatrix}$; simétrica

43. $\begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

45. $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -5 & 0 & 1 \end{pmatrix}$

47. $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & \frac{1}{5} \\ 0 & 0 & 1 \end{pmatrix}$

49. $\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

51. $\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \times \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}\begin{pmatrix} 1 & -\frac{1}{2} \\ 0 & 1 \end{pmatrix}$

53. $\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ -4 & 1 \end{pmatrix}\begin{pmatrix} 1 & -\frac{1}{2} \\ 0 & 0 \end{pmatrix}$

55. (a) $R_1 = \begin{pmatrix} 7 & -3 \\ -2 & 1 \\ 0 & 0 \end{pmatrix}$ y $R_2 = \begin{pmatrix} 8 & -2 \\ -3 & 0 \\ 1 & 1 \end{pmatrix}$

(b) $R_1\begin{pmatrix} 4 \\ -2 \end{pmatrix} = \begin{pmatrix} 34 \\ -10 \\ 0 \end{pmatrix}$ y $R_2\begin{pmatrix} 4 \\ -2 \end{pmatrix} = \begin{pmatrix} 36 \\ -12 \\ 2 \end{pmatrix}$

Capítulo 2

Problemas 2.1

1. −10 **3.** 47 **5.** 4 **7.** 56 **9.** 274

11. Sean $A = \begin{pmatrix} a_{11} & 0 & 0 & \cdots & 0 \\ 0 & a_{22} & 0 & \cdots & 0 \\ 0 & 0 & a_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{pmatrix}$

y

$B = \begin{pmatrix} b_{11} & 0 & 0 & \cdots & 0 \\ 0 & b_{22} & 0 & \cdots & 0 \\ 0 & 0 & b_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & b_{nn} \end{pmatrix}$

Entonces $\det A = a_{11}a_{22}a_{33}\cdots a_{nn}$, $\det B = b_{11}b_{22}b_{33}\cdots b_{nn}$,

$$AB = \begin{pmatrix} a_{11}b_{11} & 0 & 0 & \cdots & 0 \\ 0 & a_{22}b_{22} & 0 & \cdots & 0 \\ 0 & 0 & a_{33}b_{33} & \cdots & 0 \\ \vdots & \vdots & & & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn}b_{nn} \end{pmatrix}$$

y

$$\det AB = (a_{11}b_{11})(a_{22}b_{22}) \times (a_{33}b_{33})\cdots(a_{nn}b_{nn}) = (a_{11}a_{22}a_{33}\cdots a_{nn}) \times (b_{11}b_{22}b_{33}\cdots b_{nn}) = \det A \det B.$$

13. Casi cualquier ejemplo funcionará. Por ejemplo,

$\det\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1$, pero

$\det\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \det\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = 0 + 0 \neq 1.$

Como otro ejemplo, sean $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ y $B = \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix}$; entonces $(A+B) = \begin{pmatrix} 6 & 8 \\ 10 & 12 \end{pmatrix}$, $\det A = -2$, $\det B = -2$, y $\det(A+B) = -8 \neq \det A + \det B$.

15. Sea $A = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$.

Entonces, desarrollando continuamente en el primer renglón, se obtiene

$$\det A = a_{11}\begin{vmatrix} a_{22} & 0 & \cdots & 0 \\ a_{32} & a_{33} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ a_{n2} & a_{n3} & \cdots & a_{nn} \end{vmatrix}$$

$$= a_{11}a_{22}\begin{vmatrix} a_{33} & 0 & \cdots & 0 \\ a_{43} & a_{44} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ a_{n3} & a_{n4} & \cdots & a_{nn} \end{vmatrix}$$

$$= \cdots = a_{11}a_{22}a_{33}\cdots a_{n-2}\begin{vmatrix} a_{n-1,n-1} & 0 \\ a_{n,n-1} & a_{nn} \end{vmatrix}$$

$$= a_{11}a_{22}a_{33}\cdots a_{n-2,n-2}a_{n-1,n-1}a_{nn}.$$

17. Sean $\mathbf{u}_1 = \begin{pmatrix} u_{11} \\ u_{12} \end{pmatrix}$, $\mathbf{u}_2 = \begin{pmatrix} u_{21} \\ u_{22} \end{pmatrix}$ y $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$. Entonces

$$\mathbf{v}_1 = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\begin{pmatrix} u_{11} \\ u_{12} \end{pmatrix} = \begin{pmatrix} a_{11}u_{11} + a_{12}u_{12} \\ a_{21}u_{11} + a_{22}u_{12} \end{pmatrix},$$

$$\mathbf{v}_2 = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\begin{pmatrix} u_{21} \\ u_{22} \end{pmatrix} = \begin{pmatrix} a_{11}u_{21} + a_{12}u_{22} \\ a_{21}u_{21} + a_{22}u_{22} \end{pmatrix}.$$

Sea $A_\mathbf{v}$ el área generada por $\mathbf{v}_1$ y $\mathbf{v}_2$. Entonces, aplicando el resultado del Problema 16, resulta

$$A_\mathbf{v} = |(a_{11}u_{11} + a_{12}u_{12}) \times (a_{21}u_{21} + a_{22}u_{22}) - (a_{11}u_{21} + a_{12}u_{22}) \times (a_{21}u_{11} + a_{22}u_{12})|$$

$$= |a_{21}a_{11}u_{21}u_{11} + a_{21}a_{12}u_{21}u_{12} + a_{21}a_{11}u_{11}u_{22} + a_{22}a_{12}u_{12}u_{22} - a_{21}a_{11}u_{21}u_{11} - a_{22}a_{11}u_{21}u_{12} - a_{21}a_{12}u_{11}u_{22} - a_{22}a_{12}u_{12}u_{22}|$$

$$= |u_{11}u_{22} - u_{12}u_{21}| \times |a_{22}a_{11} - a_{21}a_{12}| = A_\mathbf{u}|\det A|$$

en donde $A_\mathbf{u}$ es el área generada por $\mathbf{u}_1$ y $\mathbf{u}_2$, empleando nuevamente el resultado del Problema 16.

Problemas 2.2

1. 28 **3.** 2 **5.** 32 **7.** −36 **9.** −260 **11.** −183 **13.** 24 **15.** −296 **17.** −138 **19.** *abcde* **21.** −8 **23.** 16 **25.** −16 **27.** −16

29. Demuestre por inducción: verdadero para $n = 2$ ya que

$$\begin{vmatrix} 1+x_1 & x_2 \\ x_1 & 1+x_2 \end{vmatrix} = (1+x_1)(1+x_2) - x_1x_2 = 1 + x_1 + x_2.$$

Suponiendo que es verdadero para $n = k$. Esto es,

$$\begin{vmatrix} 1+x_1 & x_2 & x_3 & \cdots & x_k \\ x_1 & 1+x_2 & x_3 & \cdots & x_k \\ x_1 & x_2 & 1+x_3 & \cdots & x_k \\ \vdots & \vdots & \vdots & & \vdots \\ x_1 & x_2 & x_3 & \cdots & 1+x_k \end{vmatrix} = 1 + x_1 + x_2 + \cdots + x_k.$$

Entonces, para $n = k + 1$,

$$\begin{vmatrix} 1+x_1 & x_2 & x_3 & \cdots & x_k & x_{k+1} \\ x_1 & 1+x_2 & x_3 & \cdots & x_k & x_{k+1} \\ x_1 & x_2 & 1+x_3 & \cdots & x_k & x_{k+1} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ x_1 & x_2 & x_3 & \cdots & x_k & 1+x_{k+1} \end{vmatrix}$$

(usando la Propiedad 3 en la primera columna)

$$= \begin{vmatrix} 1 & x_2 & x_3 & \cdots & x_k & x_{k+1} \\ 0 & 1+x_2 & x_3 & \cdots & x_k & x_{k+1} \\ 0 & x_2 & 1+x_3 & \cdots & x_k & x_{k+1} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & x_2 & x_3 & \cdots & x_k & 1+x_{k+1} \end{vmatrix} ① + \begin{vmatrix} x_1 & x_2 & x_3 & \cdots & x_k & x_{k+1} \\ x_1 & 1+x_2 & x_3 & \cdots & x_k & x_{k+1} \\ x_1 & x_2 & 1+x_3 & \cdots & x_k & x_{k+1} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ x_1 & x_2 & x_3 & \cdots & x_k & 1+x_{k+1} \end{vmatrix} ②$$

Pero, desarrollando det ① en su primera columna, tenemos

$$\det ① = \begin{vmatrix} 1+x_2 & x_3 & \cdots & x_k & x_{k+1} \\ x_2 & 1+x_3 & \cdots & x_k & x_{k+1} \\ \vdots & & & & \\ x_2 & x_3 & \cdots & x_k & 1+x_{k+1} \end{vmatrix} = 1 + x_2 + x_3 + \cdots + x_{k+1}$$

por el supuesto de inducción (ya que ① es un determinante de $k \times k$). Para evaluar det ②, se resta el primer renglón de todos los demás renglones:

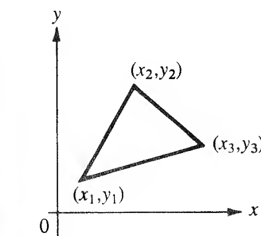
$$\det ② = \begin{vmatrix} x_1 & x_2 & x_3 & \cdots & x_k & x_{k+1} \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{vmatrix} = x_1.$$

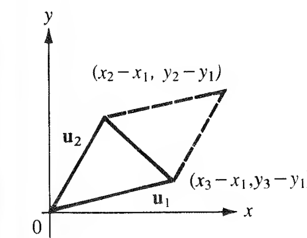
Sumando det ① y det ② se completa la demostración.

31. Si n es impar, $\det A = -\det A$, así que $2 \det A = 0$ y $\det A = 0$.

33. $\frac{1}{2}\begin{vmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{vmatrix} = \frac{1}{2}\begin{vmatrix} x_2 - x_1 & x_3 - x_1 \\ y_2 - y_1 & y_3 - y_1 \end{vmatrix}.$

Véanse las figuras que siguen:





El área A del triángulo es la mitad del área del paralelogramo generado por los vectores $\mathbf{u}_1$ y $\mathbf{u}_2$, la cual, por el resultado del Problema 3.1.16, está dada por

$$A = \pm\frac{1}{2}\begin{vmatrix} x_2 - x_1 & x_3 - x_1 \\ y_2 - y_1 & y_3 - y_1 \end{vmatrix}.$$

35. $$D_3 = \begin{vmatrix} 1 & 1 & 1 \\ a_1 & a_2 & a_3 \\ a_1^2 & a_2^2 & a_3^2 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ a_1 & a_2 - a_1 & a_3 - a_1 \\ a_1^2 & a_2^2 - a_1^2 & a_3^2 - a_1^2 \end{vmatrix}$$

$$= \begin{vmatrix} a_2 - a_1 & a_3 - a_1 \\ (a_2 + a_1)(a_2 - a_1) & (a_3 + a_1)(a_3 - a_1) \end{vmatrix}$$

$$= (a_2 - a_1)(a_3 - a_1) \times \begin{vmatrix} 1 & 1 \\ a_2 + a_1 & a_3 + a_1 \end{vmatrix}$$

$$= (a_2 - a_1)(a_3 - a_1) \times (a_3 - a_2)$$

37. **(a)** $$D_n = \begin{vmatrix} 1 & 1 & \cdots & 1 \\ a_1 & a_2 & \cdots & a_n \\ a_1^2 & a_2^2 & \cdots & a_n^2 \\ \vdots & & & \\ a_1^{n-1} & a_1^{n-1} & \cdots & a_n^{n-1} \end{vmatrix}$$

(b) Se demuestra esto por inducción. El resultado es verdadero para $n = 3$ por el resultado del Problema 35. Se supone verdadero para $n = k$. Ahora bien

$$D_{k+1} = \begin{vmatrix} 1 & 1 & \cdots & 1 & 1 \\ a_1 & a_2 & \cdots & a_k & a_{k+1} \\ a_1^2 & a_2^2 & \cdots & a_k^2 & a_{k+1}^2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_1^{k-1} & a_2^{k-1} & \cdots & a_k^{k-1} & a_{k+1}^{k-1} \\ a_1^k & a_2^k & \cdots & a_k^k & a_{k+1}^k \end{vmatrix}$$

Restamos la primera columna de cada una de las otras k columnas:

$D_{k+1} =$

$$\begin{vmatrix} 1 & 0 & \cdots & 0 & 0 \\ a_1 & a_2 - a_1 & \cdots & a_k - a_1 & a_{k+1} - a_1 \\ a_1^2 & a_2^2 - a_1^2 & \cdots & a_k^2 - a_1^2 & a_{k+1}^2 - a_1^2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_1^{k-1} & a_2^{k-1} - a_1^{k-1} & \cdots & a_k^{k-1} - a_1^{k-1} & a_{k+1}^{k-1} - a_1^{k-1} \\ a_1^k & a_2^k - a_1^k & \cdots & a_k^k - a_1^k & a_{k+1}^k - a_1^k \end{vmatrix}$$

$$= \begin{vmatrix} a_2 - a_1 & a_3 - a_1 & \cdots & a_k - a_1 & a_{k+1} - a_1 \\ a_2^2 - a_1^2 & a_3^2 - a_1^2 & \cdots & a_k^2 - a_1^2 & a_{k+1}^2 - a_1^2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_2^{k-1} - a_1^{k-1} & a_3^{k-1} - a_1^{k-1} & \cdots & a_k^{k-1} - a_1^{k-1} & a_{k+1}^{k-1} - a_1^{k-1} \\ a_2^k - a_1^k & a_3^k - a_1^k & \cdots & a_k^k - a_1^k & a_{k+1}^k - a_1^k \end{vmatrix}.$$

Ahora $a_2^k - a_1^k = (a_2 - a_1) \times (a_2^{k-1} + a_2^{k-2}a_1 + a_2^{k-3}a_1^2 + \cdots + a_2^2 a_1^{k-3} + a_2 a_1^{k-2} + a_1^{k-1})$, y $a_2^{k-1} - a_1^{k-1} = (a_2 - a_1) \times (a_2^{k-2} + a_2^{k-3}a_1 + \cdots + a_2^2 a_1^{k-4} + a_2 a_1^{k-3} + a_1^{k-2})$. Nótese que si los términos en el segundo factor de la última expresión se multiplican por a_1, y luego se restan del segundo factor de $a_2^k - a_1^k$, solamente queda el término a_2^{k-1}. De esta manera, **(i)** se desarrolla el último determinante obtenido anteriormente en el primer renglón, **(ii)** se factoriza $a_{j-1} - a_1$ en la columna j para $1 \le j \le k$, y **(iii)** se multiplica el renglón ℓ por a_1 y se le resta del renglón $(\ell + 1)$, para $\ell = k - 1, k - 2, \ldots, 3, 2$ en sucesión. Esto da

$D_{k+1} = (a_2 - a_1) \times (a_3 - a_1) \cdots (a_{k+1} - a_1)$

$$\times \begin{vmatrix} 1 & 1 & \cdots & 1 & 1 \\ a_2 & a_3 & \cdots & a_k & a_{k+1} \\ a_2^2 & a_3^2 & \cdots & a_k^2 & a_{k+1}^2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_2^{k-1} & a_3^{k-1} & \cdots & a_k^{k-1} & a_{k+1}^{k-1} \\ a_2^k & a_3^k & \cdots & a_k^k & a_{k+1}^k \end{vmatrix}$$

$$= \prod_{j=2}^{k+1} (a_j - a_1) \prod_{\substack{i=2 \\ j>i}}^{k+1} (a_j - a_i)$$

(de la hipótesis de inducción ya que el último determinante es

$$k \times k) = \prod_{\substack{i=1 \\ j>i}}^{k+1} (a_j - a_i);$$

esto completa la demostración.

39. (a) $A^2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$; $k = 2$

(b) $A^3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$; $k = 3$

41. $\det A^2 = \det A \det A = \det A$. Si $\det A \neq 0$, entonces $\det A = 1$. La respuesta es 0 o bien 1.

Problemas 2.3

1. $a_{1k} A_{1k}$ es el único término del desarrollo en la primera columna de A que contiene la componente a_{1k}. Pero

$$a_{1k} A_{1k} = a_{1k}(-1)^{1+k}|M_{1k}|.$$

Si se desarrolla $|M_{1k}|$ respecto a su columna l para $l \neq k$, el término en el desarrollo toma la forma a_{il} (cofactor de a_{il} en M_{1k}). Pero esta es la única aparición de a_{il} en el desarrollo de M_{1k}, puesto que los otros términos tienen la forma a_{jl} (cofactor de a_{jl} en M_{1k}), que suprime la columna correspondiente a la columna l de A, y a_{il} está en la columna l. Por tanto, $a_{1k} A_{1k} = (-1)^{1+k} a_{1k} a_{il} \cdot$ (cofactor de a_{il} en M_{1k}).

3. Desarrolle $|A|$ respecto a su columna k. Un término es $a_{ik} A_{ik}$ y ésta es la única aparición de a_{ij} en el desarrollo de $|A|$. Ahora $A_{ik} = (-1)^{i+k}|M_{ik}|$, y si esto se desarrolla en la columna l (para $l \neq k$), el único término en el desarrollo que contiene a a_{jl} es $a_{jl} \cdot$ (cofactor de a_{jl} en M_{ik}) por la misma razón que en el Problema 1. Así, la única aparición de $a_{ij} a_{jl}$ es $(-1)^{i+k} a_{ik} a_{jl} \cdot$ (cofactor de a_{jl} en M_{ik}).

5. -6

7. EB es la matriz obtenida multiplicando por c el renglón i de B y sumándolo al renglón j. Por la Propiedad 7,

$$\det EB = \det B =$$

$$1 \det B = \det E \det B,$$

puesto que, de (16), $\det E = 1$.

Problemas 2.4

1. $\begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{4} & \frac{3}{4} \end{pmatrix}$ **3.** $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

5. $\begin{pmatrix} \frac{1}{3} & -\frac{1}{4} & -\frac{1}{6} \\ 0 & \frac{1}{4} & \frac{1}{2} \\ 0 & \frac{1}{4} & -\frac{1}{2} \end{pmatrix}$

7. $\begin{pmatrix} 0 & 1 & -1 \\ 2 & -2 & -1 \\ -1 & 1 & 1 \end{pmatrix}$

9. no es invertible

11. $\begin{pmatrix} \frac{7}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{2}{3} \\ \frac{4}{9} & -\frac{1}{9} & -\frac{4}{9} & \frac{1}{9} \\ -\frac{1}{9} & -\frac{2}{9} & \frac{1}{9} & \frac{2}{9} \\ -\frac{5}{3} & \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \end{pmatrix}$

13. Se deduce del hecho de que $\det A^t = \det A$.

15. $A^{-1} = \begin{pmatrix} \frac{1}{14} & \frac{1}{14} & \frac{9}{28} \\ -\frac{5}{7} & \frac{2}{7} & -\frac{3}{14} \\ \frac{1}{14} & \frac{1}{14} & -\frac{5}{28} \end{pmatrix}$,

$\det A = -28$, $\det A^{-1} = -\frac{1}{28}$

17. no tiene inversa si α es cualquier número real

Problemas 2.5

1. $x_1 = -5, x_2 = 3$

3. $x_1 = 2, x_2 = 5, x_3 = -3$

5. $x_1 = \frac{45}{13}$, $x_2 = -\frac{11}{13}$, $x_3 = \frac{23}{13}$

7. $x_1 = \frac{3}{2}$, $x_2 = \frac{3}{2}$, $x_3 = \frac{1}{2}$

9. $x_1 = \frac{21}{29}$, $x_2 = \frac{171}{29}$, $x_3 = -\frac{284}{29}$, $x_4 = -\frac{182}{29}$

Ejercicios de repaso • Capítulo 2

1. -4 **3.** 24 **5.** 60 **7.** 34

9. $\begin{pmatrix} -\frac{1}{11} & \frac{4}{11} \\ \frac{2}{11} & \frac{3}{11} \end{pmatrix}$

11. no es invertible

13. $\begin{pmatrix} \frac{1}{11} & \frac{1}{11} & 0 & \frac{3}{11} \\ \frac{9}{11} & -\frac{2}{11} & 0 & -\frac{6}{11} \\ \frac{3}{11} & \frac{3}{11} & 0 & -\frac{2}{11} \\ \frac{1}{22} & \frac{1}{22} & -\frac{1}{2} & \frac{1}{22} \end{pmatrix}$

15. $x_1 = \frac{11}{7}$, $x_2 = \frac{1}{7}$

17. $x_1 = \frac{1}{4}$, $x_2 = \frac{5}{4}$, $x_3 = -\frac{3}{4}$

Capítulo 3

Problemas 3.1

1. $|\mathbf{v}| = 4\sqrt{2}$, $\theta = \pi/4$

3. $|\mathbf{v}| = 4\sqrt{2}$, $\theta = 7\pi/4$

5. $|\mathbf{v}| = 2$, $\theta = \pi/6$

7. $|\mathbf{v}| = 2$, $\theta = 2\pi/3$

9. $|\mathbf{v}| = 2$, $\theta = 4\pi/3$

11. $|\mathbf{v}| = \sqrt{89}$, $\theta = \pi + \tan^{-1}(-\frac{8}{5}) \approx 2.13$ (en el segundo cuadrante)

13. (a) $(6, 9)$ (b) $(-3, 7)$ (c) $(-7, 1)$ (d) $(39, -22)$

15. $|\mathbf{i}| = |(1, 0)| = \sqrt{1^2 + 0^2} = \sqrt{1} = 1$;

$|\mathbf{j}| = |(0, 1)| = \sqrt{0^2 + 1^2} = \sqrt{1} = 1$

17. $|\mathbf{u}| =$

$$\sqrt{\left(\frac{a}{\sqrt{a^2 + b^2}}\right)^2 + \left(\frac{b}{\sqrt{a^2 + b^2}}\right)^2}$$

$$= \sqrt{\frac{a^2}{a^2 + b^2} + \frac{b^2}{a^2 + b^2}} = 1.$$

Dirección de $|\mathbf{u}|$ =

$$\tan^{-1} \frac{\frac{b}{\sqrt{a^2 + b^2}}}{\frac{a}{\sqrt{a^2 + b^2}}} =$$

$$\tan^{-1}\left(\frac{b}{a}\right) = \text{dirección de } \mathbf{v}.$$

19. $(1/\sqrt{2})\mathbf{i} - (1/\sqrt{2})\mathbf{j}$

21. $(1/\sqrt{2})\mathbf{i} + (1/\sqrt{2})\mathbf{j}$ si $a > 0$; $-(1/\sqrt{2})\mathbf{i} - (1/\sqrt{2})\mathbf{j}$ si $a < 0$

23. sen $\theta = -3/\sqrt{13}$, cos $\theta = 2/\sqrt{13}$

25. $-(1/\sqrt{2})\mathbf{i} - (1/\sqrt{2})\mathbf{j}$

27. $\frac{3}{5}\mathbf{i} - \frac{4}{5}\mathbf{j}$

29. (a) $(1/\sqrt{2})\mathbf{i} - (1/\sqrt{2})\mathbf{j}$
(b) $(7/\sqrt{193})\mathbf{i} - (12/\sqrt{193})\mathbf{j}$
(c) $-(2/\sqrt{53})\mathbf{i} + (7/\sqrt{53})\mathbf{j}$

31. $\overrightarrow{PQ}$ es una representación de $(c + a - c)\mathbf{i} + (d + b - d)\mathbf{j} = a\mathbf{i} + b\mathbf{j}$. Por consiguiente, $\overrightarrow{PQ}$ y (a, b) son representaciones del mismo vector.

33. $4\mathbf{i} + 4\sqrt{3}\mathbf{j}$ **35.** $-3\mathbf{i} + 3\sqrt{3}\mathbf{j}$

37. **(i)** Supóngase que $\mathbf{u} = \alpha\mathbf{v}$, en donde $\alpha > 0$. Entonces $|\mathbf{u} + \mathbf{v}| = |\alpha\mathbf{v} + \mathbf{v}| = |(\alpha + 1)\mathbf{v}| = |\alpha + 1|\ |\mathbf{v}| = (\alpha + 1)|\mathbf{v}|$ (puesto que $\alpha + 1 > 0$) $= \alpha|\mathbf{v}| + |\mathbf{v}| = |\alpha\mathbf{v}| + |\mathbf{v}| = |\mathbf{u}| + |\mathbf{v}|$.

(ii) Recíprocamente, considérese que $\mathbf{u} = (a, b)$, $\mathbf{v} = (c, d)$, y $|\mathbf{u} + \mathbf{v}| = |\mathbf{u}| + |\mathbf{v}|$. Entonces $\mathbf{u} + \mathbf{v} = (a + c, b + d)$, y $|\mathbf{u} + \mathbf{v}|^2 = (|\mathbf{u}| + |\mathbf{v}|)^2 = |\mathbf{u}|^2 + 2|\mathbf{u}||\mathbf{v}| + |\mathbf{v}|^2$, lo cual implica que

$$(a + c)^2 + (b + d)^2 = a^2 + b^2 + 2\sqrt{(a^2 + b^2)(c^2 + d^2)} + c^2 + d^2$$

y, después de multiplicar y eliminar términos semejantes, resulta $ac + bd = \sqrt{a^2c^2 + a^2d^2 + b^2c^2 + b^2d^2}$. De modo que, elevando al cuadrado ambos miembros y de nuevo cancelando términos semejantes, queda $2abcd = a^2d^2 + b^2c^2$, o bien $(ad - bc)^2 = a^2d^2 - 2abcd + b^2c^2 = 0$ de modo que $ad = bc$. Si $d \neq 0$, entonces $a = \frac{b}{d}c, b = \frac{b}{d}d$, y $|\mathbf{u}| = \left|\frac{b}{d}\right||\mathbf{v}|$. Entonces $\left|\frac{b + d}{d}\right||\mathbf{v}| = \left|\frac{b + d}{d}(c, d)\right| = \left|\left(\frac{b}{d}c + c, \frac{b}{d}d + d\right)\right| = |(a + c, b + d)| = |\mathbf{u} + \mathbf{v}| = |\mathbf{u}| + |\mathbf{v}| = \left|\frac{b}{d}\mathbf{v}\right| + |\mathbf{v}| = \left(\left|\frac{b}{d}\right| + 1\right)|\mathbf{v}| = \frac{|b| + |d|}{|d|}|\mathbf{v}|$. Por lo tanto $\left|\frac{b + d}{d}\right| = \frac{|b| + |d|}{|d|}$ de manera que $|b + d| = |b| + |d|$. Como b y d son números reales, esto implica que b y d tienen el mismo signo y así $\frac{b}{d}$ es positivo. Consecuentemente si $\alpha = \left|\frac{b}{d}\right| = \frac{b}{d}$, entonces $\mathbf{u} = \alpha\mathbf{v}$. Si $d = 0$, entonces $c \neq 0$ y $\mathbf{u} = \frac{a}{c}\mathbf{v}$ por el razonamiento anterior, en donde $\frac{a}{c} > 0$. Por

consiguiente, $\mathbf{u}$ es un múltiplo escalar positivo de $\mathbf{v}$.

Problemas 3.2

1. $0; 0$ **3.** $0; 0$ **5.** $20; \frac{20}{29}$

7. $-22; -22/5\sqrt{53}$ **11.** paralelos

9. $\mathbf{u}\cdot\mathbf{v} = \alpha\beta - \beta\alpha = 0$

13. ni una ni otra cosa

15. ortogonales

17. (a) $-\frac{3}{4}$ (b) $\frac{4}{3}$ (c) $\frac{1}{7}$
(d) $(-96 \pm \sqrt{7500})/78 \approx -0.12, -2.34$

19. Si $\mathbf{u}$ y $\mathbf{v}$ tienen direcciones opuestas, entonces $\theta_{\mathbf{u}} = \theta_{\mathbf{v}} + \pi$. En consecuencia, $\cos\theta_{\mathbf{u}} = \cos(\theta_{\mathbf{v}} + \pi) = -\cos\theta_{\mathbf{v}}$. Esto implica que

$$\cos\theta_{\mathbf{u}} = \tfrac{3}{5} = -\frac{1}{\sqrt{1+\alpha^2}} =$$

$-\cos\theta_{\mathbf{v}}$, o bien $\sqrt{1+\alpha^2} = -\frac{5}{3} < 0$, lo cual es imposible dado que $\sqrt{1+\alpha^2} = |\mathbf{v}| \geq 0$.

21. $\frac{3}{2}\mathbf{i} + \frac{3}{2}\mathbf{j}$ **23.** $\mathbf{0}$ **25.** $-\frac{2}{13}\mathbf{i} + \frac{3}{13}\mathbf{j}$

27. $[(\alpha+\beta)/2]\mathbf{i} + [(\alpha+\beta)/2]\mathbf{j}$

29. $[(\alpha-\beta)/2]\mathbf{i} + [(\alpha-\beta)/2]\mathbf{j}$

31. $a_1a_2 + b_1b_2 \geq 0$

33. $\text{Proy}_{\overrightarrow{PQ}}\, \overrightarrow{RS} = \frac{51}{25}\mathbf{i} + \frac{68}{25}\mathbf{j}$;
$\text{Proy}_{\overrightarrow{RS}}\, \overrightarrow{PQ} = -\frac{17}{26}\mathbf{i} + \frac{85}{26}\mathbf{j}$

35. **(i)** Si $\mathbf{u} = \alpha\mathbf{v}$, entonces $\mathbf{u}\cdot\mathbf{v} = \alpha\mathbf{v}\cdot\mathbf{v} = \alpha|\mathbf{v}|^2$, $|\mathbf{u}| = |\alpha||\mathbf{v}|^2$ de modo que

$$\cos\phi = \frac{\alpha|\mathbf{v}|^2}{|\alpha||\mathbf{v}||\mathbf{v}|} = \frac{\alpha}{|\alpha|} = \pm 1.$$

(ii) Supóngase que $\mathbf{u}$ y $\mathbf{v}$ son paralelos. Entonces, si $\mathbf{u} = (a, b)$ y $\mathbf{v} = (c, d)$, se tiene

$$1 = \cos^2\phi = \frac{|\mathbf{u}\cdot\mathbf{v}|^2}{|\mathbf{u}|^2|\mathbf{v}|^2} = \frac{(ac+bd)^2}{(a^2+b^2)(c^2+d^2)}.$$

Multiplicando y simplificando, obtenemos $0 = a^2d^2 - 2abcd + b^2c^2$. Por tanto $ad = bc$. Si $a \neq 0$, entonces

$$d = \left(\frac{c}{a}\right)b \text{ y } c = \left(\frac{c}{a}\right)a$$

de manera quc $\mathbf{v} = \left(\frac{c}{a}\right)\mathbf{u}$.

Si $a = 0$, entonces $b \neq 0$ y $\mathbf{v} = \frac{d}{b}\mathbf{u}$.

37. La recta $ax + by + c = 0$ tiene pendiente $-\frac{a}{b}$. Un vector paralelo a la recta es $\mathbf{u} = \mathbf{i} - \frac{a}{b}\mathbf{j}$, y

$$\mathbf{u}\cdot\mathbf{v} = 1\cdot a - \frac{a}{b}\cdot b = 0.$$

39. $52/5\sqrt{113} \approx 0.9783$; $61/\sqrt{34}\sqrt{113} \approx 0.9841$; $-27/5\sqrt{34} \approx -0.9261$

41. Si $a_1 = a_2 = 0$ o bien $b_1 = b_2 = 0$, ambos miembros de la desigualdad son iguales a cero. Si por lo menos uno de a_1 y $a_2 \neq 0$ y al menos uno de b_1 y $b_2 \neq 0$, sean $\mathbf{u} = a_1\mathbf{i} + a_2\mathbf{j}$, $\mathbf{v} = b_1\mathbf{i} + b_2\mathbf{j}$. Entonces $\mathbf{u} \neq 0$, $\mathbf{v} \neq 0$, $|\mathbf{u}||\mathbf{v}| \neq 0$,

y $\left|\dfrac{|\mathbf{u}\cdot\mathbf{v}|}{|\mathbf{u}||\mathbf{v}|}\right| = |\cos\phi| \leq 1$.

Por consiguiente, $|a_1b_1 + a_2b_2| = |\mathbf{u}\cdot\mathbf{v}| \leq |\mathbf{u}||\mathbf{v}| = \sqrt{a_1^2 + a_2^2}\sqrt{b_1^2 + b_2^2}$.

La igualdad se verifica cuando $|\cos\phi| = 1$, lo cual es cierto si y sólo si $\mathbf{u}$ y $\mathbf{v}$ son paralelos.

43. $\sqrt{5}$

45. Sea $A = \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}$. Entonces $A^t = \begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix}$. Sean $\mathbf{u} = (a_1, a_2)$ y $\mathbf{v} = (b_1, b_2)$. Entonces $\mathbf{u}\cdot\mathbf{v} = 0$, $|\mathbf{u}| = 1$ y $|\mathbf{v}| = 1$.

$$A^tA = \begin{pmatrix} a_1^2 + b_1^2 & a_1a_2 + b_1b_2 \\ a_2a_1 + b_2b_1 & a_2^2 + b_2^2 \end{pmatrix} = \begin{pmatrix} |\mathbf{u}|^2 & \mathbf{u}\cdot\mathbf{v} \\ \mathbf{u}\cdot\mathbf{v} & |\mathbf{v}|^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

De igual manera $AA^t = I$. Consecuentemente A es invertible y $A^{-1} = A^t$.

Problemas 3.3

1. $\sqrt{40}$ **3.** 6

5. $3; -1, 0, 0$

7. $\sqrt{5}$; $1/\sqrt{5}, 0, 2/\sqrt{5}$

9. $\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3}, -1/\sqrt{3}$

11. $\sqrt{3}, 1/\sqrt{3}, -1/\sqrt{3}, -1/\sqrt{3}$

13. $\sqrt{3}, -1/\sqrt{3}, -1/\sqrt{3}, 1/\sqrt{3}$

15. $\sqrt{78}$; $2/\sqrt{78}, 5/\sqrt{78}, -7/\sqrt{78}$

17. $\sqrt{29}$; $-2/\sqrt{29}, -3/\sqrt{29}, -4/\sqrt{29}$

19. $4\sqrt{3}\mathbf{i} + 4\sqrt{3}\mathbf{j} + 4\sqrt{3}\mathbf{k}$

21. $(1/\sqrt{26})\mathbf{i} - (3/\sqrt{26})\mathbf{j} + (4/\sqrt{26})\mathbf{k}$

23. $R = (-3, y, z)$, y, z arbitrarios; este conjunto de puntos constituye un plano paralelo al plano yz.

25. $\left|\dfrac{\mathbf{u}\cdot\mathbf{v}}{|\mathbf{u}||\mathbf{v}|}\right| = |\cos\phi| \leq 1$. Por lo tanto $|\mathbf{u}\cdot\mathbf{v}| \leq |\mathbf{u}||\mathbf{v}|$. Entonces

$$\begin{aligned}|\mathbf{u}+\mathbf{v}|^2 &= (\mathbf{u}+\mathbf{v})\cdot(\mathbf{u}+\mathbf{v}) \\ &= |\mathbf{u}|^2 + 2\mathbf{u}\cdot\mathbf{v} + |\mathbf{v}|^2 \\ &\leq |\mathbf{u}|^2 + 2|\mathbf{u}||\mathbf{v}| + |\mathbf{v}|^2 \\ &= (|\mathbf{u}| + |\mathbf{v}|)^2.\end{aligned}$$

27. $-6\mathbf{j} + 9\mathbf{k}$ **29.** $8\mathbf{i} - 14\mathbf{j} + 9\mathbf{k}$

31. $16\mathbf{i} + 29\mathbf{j} + 42\mathbf{k}$ **33.** $\sqrt{59}$

35. $\cos^{-1}(35/\sqrt{29}\sqrt{59}) \approx \cos^{-1}(0.8461) \approx 0.5621 \approx 32.21^{0}$

37. $\frac{25}{29}\mathbf{u} = \frac{50}{29}\mathbf{i} - \frac{75}{29}\mathbf{j} + \frac{100}{29}\mathbf{k}$

39. Como los segmentos de recta PS y SR son perpendiculares (en la Figura 3.28), el triángulo PSR es un triángulo rectángulo y

$$\overline{PR}^2 = \overline{PS}^2 + \overline{SR}^2 \quad \text{(i)}$$

Pero PRQ es también un triángulo rectángulo de modo que

$$\overline{PQ}^2 = \overline{PR}^2 + \overline{RQ}^2 \quad \text{(ii)}$$

Así que, combinando (i) y (ii), resulta

$$\overline{PQ}^2 = \overline{PS}^2 + \overline{SR}^2 + \overline{RQ}^2 \quad \text{(iii)}$$

Puesto que las coordenadas x y z de P y S son iguales,

$$\overline{PS}^2 = (y_2 - y_1)^2 \quad \text{(iv)}$$

Análogamente,

$$\overline{RS}^2 = (x_2 - x_1)^2 \quad \text{(v)}$$

y $$\overline{RQ}^2 = (z_2 - z_1)^2 \quad \text{(vi)}$$

Por lo tanto, usando (iv), (v) y (vi) en (iii) da

$$\overline{PQ}^2 = (x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2$$

41. **(i)** Si $\mathbf{v} = \alpha\mathbf{u}$, entonces

$$\cos\phi = \frac{\mathbf{u}\cdot\mathbf{v}}{|\mathbf{u}||\mathbf{v}|} = \frac{\alpha|\mathbf{u}|^2}{|\alpha||\mathbf{u}|^2} = \pm 1.$$

Si $\mathbf{u}$ y $\mathbf{v}$ son paralelos, entonces $\dfrac{\mathbf{u}}{|\mathbf{u}|} = \pm\dfrac{\mathbf{v}}{|\mathbf{v}|}$ y así $\mathbf{v} = \pm\dfrac{|\mathbf{v}|}{|\mathbf{u}|}\mathbf{u} = \alpha\mathbf{u}$.

(ii) Si $\mathbf{u}\cdot\mathbf{v} = 0$, entonces $\cos\phi = 0$ y $\phi = \frac{\pi}{2}$. Si $\phi = \frac{\pi}{2}$, entonces $\mathbf{u}\cdot\mathbf{v} = |\mathbf{u}||\mathbf{v}|\cos\phi = 0$.

Problemas 3.4

1. $-6\mathbf{i} - 3\mathbf{j}$ **3.** $-\mathbf{i} - \mathbf{j} + \mathbf{k}$

5. $12\mathbf{i} + 8\mathbf{j} - 21\mathbf{k}$

7. $(bc - ad)\mathbf{j}$

9. $-5\mathbf{i} - \mathbf{j} + 7\mathbf{k}$ **11.** $\mathbf{0}$

13. $42\mathbf{i} + 6\mathbf{j}$

15. $-9\mathbf{i} + 39\mathbf{j} + 61\mathbf{k}$

17. $-4i + 8\mathbf{k}$ **19.** $\mathbf{0}$

21. $\pm[-(9/\sqrt{181})\mathbf{i} - (6/\sqrt{181})\mathbf{j} + (8/\sqrt{181})\mathbf{k}]$

23. $\sqrt{30}/\sqrt{6}\sqrt{29} \approx 0.415$

25. $5\sqrt{5}$ **27.** $\sqrt{523}$

29. $\sqrt{a^2b^2 + a^2c^2 + b^2c^2}$

31. Sean $\mathbf{u} = a_1\mathbf{i} + b_1\mathbf{j} + c_1\mathbf{k}$ y $\mathbf{v} = a_2\mathbf{i} + b_2\mathbf{j} + c_2\mathbf{k}$. Entonces $\mathbf{u}\times\mathbf{v} = (b_1c_2 - c_1b_2)\mathbf{i} + (c_1a_2 - a_1c_2)\mathbf{j} + (a_1b_2 - b_1a_2)\mathbf{k}$ de manera que $|\mathbf{u}\times\mathbf{v}|^2 = (b_1c_2 - c_1b_2)^2 + (c_1a_2 - a_1c_2)^2 + (a_1b_2 - b_1a_2)^2 = b_1^2c_2^2 - 2b_1c_2c_1b_2 + c_1^2b_2^2 + c_1^2a_2^2 - 2c_1a_2a_1c_2 + a_1^2c_2^2 + a_1^2b_2^2 - 2a_1b_2b_1a_2 + b_1^2a_2^2$. Esto equivale a $|\mathbf{u}|^2|\mathbf{v}|^2 - (\mathbf{u}\cdot\mathbf{v})^2 = (a_1^2 + b_1^2 + c_1^2)(a_2^2 + b_2^2 + c_2^2) - (a_1a_2 + b_1b_2 + c_1c_2)^2$.

33. Sean $u = a_1\mathbf{i} + b_1\mathbf{j} + c_1\mathbf{k}$, $\mathbf{v} = a_2\mathbf{i} + b_2\mathbf{j} + c_2\mathbf{k}$, y $\mathbf{w} = a_3\mathbf{i} + b_3\mathbf{j} + c_3\mathbf{k}$
Entonces

$$\begin{aligned}(\mathbf{u}\times\mathbf{v})\cdot\mathbf{w} &= [(b_1c_2 - c_1b_2)\mathbf{i} + (c_1a_2 - a_1c_2)\mathbf{j} + (a_1b_2 - b_1a_2)\mathbf{k}]\cdot[a_3\mathbf{i} + b_3\mathbf{j} + c_3\mathbf{k}] \\ &= b_1c_2a_3 - c_1b_2a_3 + c_1a_2b_3 - a_1c_2b_3 + a_1b_2c_3 - b_1a_2c_3\end{aligned}$$

y

$$\begin{aligned}\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w}) &= [a_1\mathbf{i} + b_1\mathbf{j} + c_1\mathbf{k}]\cdot[(b_2c_3 - c_2b_3)\mathbf{i} + (c_2a_3 - a_2c_3)\mathbf{j} + (a_2b_3 - b_2a_3)\mathbf{k}] \\ &= a_1b_2c_3 - a_1c_2b_3 + b_1c_2a_3 - b_1a_2c_3 + c_1a_2b_3 - c_1b_2a_3\end{aligned}$$

35. Si $\mathbf{u}$ y $\mathbf{v}$ son paralelos y ninguno es $\mathbf{0}$, entonces $\mathbf{v} = t\mathbf{u}$ para una constante t. Entonces, si $\mathbf{u} = a\mathbf{i} + b\mathbf{j} + c\mathbf{k}$,

$$\mathbf{u}\times\mathbf{v} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a & b & c \\ ta & tb & tc \end{vmatrix}$$

$= 0$ por la Propiedad 6 en la página 107.

37. El volumen de un paralelepípedo se obtiene multiplicando la base por la altura. La magnitud de $\mathbf{u}\times\mathbf{v}$ es el área de la base: Su dirección es perpendicular a la base. La altura se mide a lo largo de $\mathbf{u}\times\mathbf{v}$. La altura h es la proyección del tercer vector, $\mathbf{w}$, sobre $\mathbf{u}\times\mathbf{v}$. $h = |\text{Proy}_{\mathbf{u}\times\mathbf{v}}\mathbf{w}| = \dfrac{\mathbf{w}\cdot(\mathbf{u}\times\mathbf{v})}{|\mathbf{u}\times\mathbf{v}|}$; volumen $= (h)(\text{base}) = \dfrac{\mathbf{w}\cdot(\mathbf{u}\times\mathbf{v})}{|\mathbf{u}\times\mathbf{v}|}|\mathbf{u}\times\mathbf{v}| = (\mathbf{u}\times\mathbf{v})\cdot\mathbf{w}$ si el producto escalar es positivo, $|(\mathbf{u}\times\mathbf{v})\cdot\mathbf{w}|$ de otra manera.

39. 23

41. Este problema dependerá en alto grado de la Propiedad 3 de los determinantes, la cual establece que si la i-ésima columna (o renglón) de un determinante consiste en un par de elementos, el determinante puede volver a escribirse como una suma de dos determinantes cuyas columnas (o renglones) son idénticas, excepto por la i-ésima columna (o renglón). El primer determinante contiene uno de los elementos del par, en tanto que el otro miembro de cada par de elementos de la i-ésima columna (o renglón) aparece en el segundo determinante. Nótese también que el volumen generado por $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ está dado por

$$\text{Volumen} = \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix}$$

Sea

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix},$$

$\mathbf{u}_1 = A\mathbf{u}$, $\mathbf{v}_1 = A\mathbf{v}$, $\mathbf{w}_1 = A\mathbf{w}$.

Por consiguiente

$$\mathbf{u}_1 = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}\begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = \begin{pmatrix} a_{11}u_1 + a_{12}u_2 + a_{13}u_3 \\ a_{21}u_1 + a_{22}u_2 + a_{23}u_3 \\ a_{31}u_1 + a_{32}u_2 + a_{33}u_3 \end{pmatrix}$$

Análogamente,

$$\mathbf{v}_1 = \begin{pmatrix} a_{11}v_1 + a_{12}v_2 + a_{13}v_3 \\ a_{21}v_1 + a_{22}v_2 + a_{23}v_3 \\ a_{31}v_1 + a_{32}v_2 + a_{33}v_3 \end{pmatrix}$$

y

$$\mathbf{w}_1 = \begin{pmatrix} a_{11}w_1 + a_{12}w_2 + a_{13}w_3 \\ a_{21}w_1 + a_{22}w_2 + a_{23}w_3 \\ a_{31}w_1 + a_{32}w_2 + a_{33}w_3 \end{pmatrix}$$

Por el Problema 36, el volumen generado por $\mathbf{u}_1$, $\mathbf{v}_1$, $\mathbf{w}_1$ es:

$$V = \begin{vmatrix} a_{11}u_1 + a_{12}u_2 + a_{13}u_3 & a_{21}u_1 + a_{22}u_2 + a_{23}u_3 & a_{31}u_1 + a_{32}u_2 + a_{33}u_3 \\ a_{11}v_1 + a_{12}v_2 + a_{13}v_3 & a_{21}v_1 + a_{22}v_2 + a_{23}v_3 & a_{31}v_1 + a_{32}v_2 + a_{33}v_3 \\ a_{11}w_1 + a_{12}w_2 + a_{13}w_3 & a_{21}w_1 + a_{22}w_2 + a_{23}w_3 & a_{31}w_1 + a_{32}w_2 + a_{33}w_3 \end{vmatrix}$$

Desarrollando se puede verificar que

$$V = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix}$$

$=$ (det A)(volumen generado por **u**, **v**, **w**), si det $A \geq 0$, de otro modo utilice $-\det A$.

43. Sean $\mathbf{u} = (u_1, u_2, u_3)$, $\mathbf{v} = (v_1, v_2, v_3)$ y $\mathbf{w} = (w_1, w_2, w_3)$. Entonces

$$\mathbf{v} \times \mathbf{w} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix} = (v_2w_3 - v_3w_2,\ v_3w_1 - v_1w_3,\ v_1w_2 - v_2w_1);$$

$$\mathbf{u} \times (\mathbf{v} \times \mathbf{w}) = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ u_1 & u_2 & u_3 \\ v_2w_3 - v_3w_2 & v_3w_1 - v_1w_3 & v_1w_2 - v_2w_1 \end{vmatrix}$$

$= (-u_3v_3w_1 + u_3v_1w_3 + u_2v_1w_2 - u_2v_2w_1, u_3v_2w_3 - u_3v_3w_2 - u_1v_1w_2 + u_1v_2w_1, u_1v_3w_1 - u_1v_1w_3 - u_2v_2w_3 + u_2v_3w_2)$(*)

$(\mathbf{u}\cdot\mathbf{w})\mathbf{v} = (u_1w_1 + u_2w_2 + u_3w_3)(v_1, v_2, v_3) = (u_1v_1w_1 + u_2v_1w_2 + u_3v_1w_3, u_1v_2w_1 + u_2v_2w_2 + u_3v_2w_3, u_1v_3w_1 + u_2v_3w_2 + u_3v_3w_3)$

$-(\mathbf{u}\cdot\mathbf{v})\mathbf{w} = -(u_1v_1 + u_2v_2 + u_3w_3)(w_1, w_2, w_3) = (-u_1v_1w_1 - u_2v_2w_1 - u_3v_3w_1, -u_1v_1w_2 - u_2v_2w_2 - u_3v_3w_2, -u_1v_1w_3 - u_2v_2w_3 - u_3v_3w_3)$.

Si se suman los últimos dos vectores, obtenemos el vector (*).

Problemas 3.5

En las respuestas a los Problemas del 1 al 5 suponemos que el primer punto es P y que el segundo punto es Q. Las ecuaciones vectoriales son de la forma $\overrightarrow{QR} = \overrightarrow{QP} + t\mathbf{v}$. Sólo **v** se proporciona en las respuestas.

1. $\mathbf{v} = -\mathbf{i}+\mathbf{j}-4\mathbf{k}$; $x = 2 - t$, $y = 1 + t$, $z = 3 - 4t$; $(x-2)/(-1) = y - 1 = (z-3)/(-4)$
3. $\mathbf{v} = -\mathbf{j} - 2\mathbf{k}$; $x = -4$, $y = 1 - t$, $z = 3 - 2t$; $x = -4$ y $z = 1 + 2y$
5. $\mathbf{v} = 2\mathbf{i} - 2\mathbf{k}$; $x = 1 + 2t$, $y = 2$, $z = 3 - 2t$; $y = 2$ y $x = 4 - z$

En los Problemas del 7 al 11 ya está dado **v**.

7. $x = 2 + 2t$, $y = 2 - t$, $z = 1 - t$; $(x-2)/2 = (y-2)/(-1) = (z-1)/(-1)$
9. $x = -1$, $y = -2 - 3t$, $z = 5 + 7t$; $x = -1$ y $7y + 3z = 1$
11. $x = a + dt$, $y = b + et$, $z = c$; $(x-a)/d = (y-b)/e$ y $z = c$
13. $\mathbf{v} = 3\mathbf{i} + 6\mathbf{j} + 2\mathbf{k}$; $x = 4 + 3t$, $y = 1 + 6t$, $z = -6 + 2t$; $(x-4)/3 = (y-1)/6 = (z+6)/2$
15. El vector $\mathbf{v}_1 = a_1\mathbf{i} + b_1\mathbf{j} + c_1\mathbf{k}$ es paralelo a L_1, en tanto que el vector $\mathbf{v}_2 = a_2\mathbf{i} + b_2\mathbf{j} + c_2\mathbf{k}$ es paralelo a L_2. De modo que $L_1 \perp L_2$ si $\mathbf{v}_1 \perp \mathbf{v}_2$ o bien $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$. Pero $\mathbf{v}_1 \cdot \mathbf{v}_2 = a_1a_2 + b_1b_2 + c_1c_2$.
17. $3\mathbf{i} + 6\mathbf{j} + 9\mathbf{k} = 3(\mathbf{i} + 2\mathbf{j} + 3\mathbf{k})$ por lo que son paralelos los vectores de dirección de las rectas. Obsérvese que no coinciden puesto que, por ejemplo, el punto $(1, -3, -3)$ está en L_1 pero no en L_2.
19. Si hubieran tenido un punto en común, entonces

$$2 - t = 1 + s$$
$$1 + t = -2s$$
$$-2t = 3 + 2s$$

La solución única de las dos primeras de estas ecuaciones es $s = -2$, $t = 3$; pero este par no satisface la tercera ecuación.

21. (a) $(\sqrt{186}/3)$ $(t = \frac{1}{3})$ (b) $\sqrt{1518}/11 = \sqrt{138/11}$, $(t = -$[illegible]$)$ (c) $\sqrt{750}/6 = 5\sqrt{30}/6$ $(t = -\frac{1}{6})$
23. $(x+4)/26 = (y-7)/1 = (t-3)/37$
25. $(x-4)/(-4) = (y-6)/16 = z/24$
27. 3 29. $y = 0$ (plano xz)
31. $x + y = 3$ 33. $y + z = 5$
35. $-3x - 4y + z = 45$
37. $2x - 7y - 8z = -20$
39. $-12x - 21y + 22z = 63$
41. $2x + y = 7$ 43. coincidentes
45. no son de ninguna de esas clases
47. $(x, y, z) = (-1, -3, 0) + t(1, 2, 1)$
49. $(x, y, z) = (-11/4, 3/2, 0) + t(9, 16, 2)$
51. $13/\sqrt{69}$ 53. $19/\sqrt{35}$
55. $\cos^{-1}(9/\sqrt{3}\sqrt{29}) = \cos^{-1}(0.9649) \approx 0.2657 \approx 15.23°$
57. $\cos^{-1}|20/\sqrt{294}\sqrt{6}| = \cos^{-1}(\frac{20}{42}) \approx 1.074 \approx 61.56°$
59. $\mathbf{n} = \mathbf{v} \times \mathbf{w}$ es ortogonal a **v** y a **w**. Si $\mathbf{u}\cdot(\mathbf{v} \times \mathbf{w}) = 0$, entonces $\mathbf{u} \perp \mathbf{n}$, lo que significa que **u** está en el plano determinado por **v** y **w**.
61. coplanares; $29x - y + 11z = 0$
63. no coplanares; $\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w}) = -9$.

Ejercicios de repaso • Capítulo 3.

1. $|\mathbf{v}| = 3\sqrt{2}$, $\theta = \pi/4$
3. $|\mathbf{v}| = 4$, $\theta = 5\pi/3$
5. $|\mathbf{v}| = 12\sqrt{2}$, $\theta = 5\pi/4$
7. $2\mathbf{i} + 2\mathbf{j}$ 9. $4\mathbf{i} + 2\mathbf{j}$
11. (a) $(10, 5)$, (b) $(5, -3)$, (c) $(-31, 12)$
13. $(1/\sqrt{2})\mathbf{i} + (1/\sqrt{2})\mathbf{j}$
15. $(2/\sqrt{29})\mathbf{i} + (5/\sqrt{29})\mathbf{j}$
17. $\frac{3}{5}\mathbf{i} + \frac{4}{5}\mathbf{j}$
19. $(1/\sqrt{2})\mathbf{i} - (1/\sqrt{2})\mathbf{j}$ si $a > 0$ y $-(1/\sqrt{2})\mathbf{i} + (1/\sqrt{2})\mathbf{j}$ si $a < 0$
21. $-(5/\sqrt{29})\mathbf{i} - (2/\sqrt{29})\mathbf{j}$
23. $-(10/\sqrt{149})\mathbf{i} + (7/\sqrt{149})\mathbf{j}$
25. **j** 27. $-\frac{7}{2}\sqrt{3}\mathbf{i} + \frac{7}{2}\mathbf{j}$ 29. 0; 0
31. $-14, -14/\sqrt{5}\sqrt{41}$
33. no son de ninguna de esas clases
35. paralelos
37. paralelos 39. $7\mathbf{i} + 7\mathbf{j}$
41. $\frac{15}{13}\mathbf{i} + \frac{10}{13}\mathbf{j}$ 43. $-\frac{3}{2}\mathbf{i} - \frac{7}{2}\mathbf{j}$
45. $\text{Proy}_{\overrightarrow{RS}}\overrightarrow{PQ} = -\frac{99}{25}\mathbf{i} + \frac{132}{25}\mathbf{j}$; $\text{Proy}_{\overrightarrow{PQ}}\overrightarrow{RS} = -\frac{33}{82}\mathbf{i} - \frac{297}{82}\mathbf{j}$
47. $\sqrt{216}$
49. $\sqrt{130}; 0, 3/\sqrt{130}, 11/\sqrt{130}$
51. $\sqrt{53}; -4/\sqrt{53}, 1/\sqrt{53}, 6/\sqrt{53}$
53. $(2/\sqrt{6})\mathbf{i} - (1/\sqrt{6})\mathbf{j} + (1/\sqrt{6})\mathbf{k}$
55. $\mathbf{i} - 14\mathbf{j} + 20\mathbf{k}$
57. $\frac{26}{21}\mathbf{i} - \frac{52}{21}\mathbf{j} + \frac{13}{21}\mathbf{k}$ 59. 22
61. $\cos^{-1}(-9/\sqrt{798}) \approx 1.895 \approx 108.6°$
63. $-7\mathbf{i} - 7\mathbf{k}$
65. $-26\mathbf{i} - 8\mathbf{j} + 7\mathbf{k}$
67. $\sqrt{2065}$
69. $\overrightarrow{OR} = (-4\mathbf{i} + \mathbf{j}) + t(7\mathbf{i} - \mathbf{j} + 7\mathbf{k})$; $x = -4 + 7t$, $y = 1 - t$, $z = 7t$; $(x+4)/7 = (y-1)/(-1) = z/7$
71. $\overrightarrow{OR} = \mathbf{i} - 2\mathbf{j} - 3\mathbf{k} + t(5\mathbf{i} - 3\mathbf{j} + 2\mathbf{k})$; $x = 1 + 5t$, $y = -2 - 3t$, $z = -3 + 2t$; $(x-1)/5 = (y+2)/(-3) = (z+3)/2$
73. $\sqrt{165}/3$ 75. $x + z = -1$
77. $2x - 3y + 5z = 19$
79. $x = \frac{1}{2} - \frac{9}{2}t$, $y = \frac{7}{2} - \frac{11}{2}t$, $z = t$
81. $x = \frac{4}{3} - \frac{5}{2}t$, $y = -4 - \frac{7}{2}t$, $z = t$
83. $\cos^{-1}|-1/\sqrt{207}| = \cos^{-1} 1/\sqrt{207} \approx 1.501 \approx 86.01°$

Capítulo 4

Problemas 4.2

1. sí
3. no, (iv); también (vi) no se verifica si $\alpha < 0$
5. sí 7. sí
9. no (i), (iii), (iv), (vi) no se cumplen
11. sí 13. sí
15. no; (i), (iii), (iv), (vi) no se cumplen
17. sí 19. sí
21. Supóngase que **0** y **0**′ son identidades aditivas. Entonces, por la definición de identidad aditiva $\mathbf{0} = \mathbf{0} + \mathbf{0}'$ y $\mathbf{0}' = \mathbf{0}' + \mathbf{0} = \mathbf{0} + \mathbf{0}'$. Por tanto, $\mathbf{0} = \mathbf{0}'$.
23. Sean $-\mathbf{x}$ y $-\mathbf{x}'$ inversos aditivos de $\mathbf{x}'$. Entonces $-\mathbf{x} = -\mathbf{x} + \mathbf{0} = -\mathbf{x} + [\mathbf{x}' + (-\mathbf{x}')] = (-\mathbf{x} + \mathbf{x}') + (-\mathbf{x}')$ (por ii) $= \mathbf{0} + (-\mathbf{x}')$ (ya que $-\mathbf{x}$ es un inverso aditivo de $\mathbf{x}'$) $= -\mathbf{x}'$. En consecuencia, $-\mathbf{x} = -\mathbf{x}'$ y el inverso aditivo es único.
25. Sean y_1 y y_2 soluciones de la ecuación. Entonces

$$y_1'' + a(x)y_1' + b(x)y_1(x) = 0$$

y

$$y_2'' + a(x)y_2' + b(x)y_2(x) = 0$$

De modo que

$$(y_1 + y_2)'' + a(x)(y_1 + y_2)' + b(x)(y_1 + y_2) = +[y_1'' + a(x)y_1' + b(x)y_1] + [y_2'' + a(x)y_2' + b(x)y_2] = 0 + 0 = 0$$

y así $y_1 + y_2$ es una solución. De manera semejante $(\alpha y_1)'' + a(x)(\alpha y_1') + b(x)(\alpha y) = \alpha[y_1'' + a(x)y_1' + b(x)y_1] = \alpha \cdot 0 = 0$ así que αy_1 también es una solución. Por lo tanto son válidas las reglas de cerradura. Como $-y_1 = (-1)y$, también es una solución, se tiene el inverso aditivo. Los otros axiomas se deducen fácilmente.

Problemas 4.3

1. no; porque $\alpha(x, y) \notin H$ si $\alpha < 0$
3. sí 5. sí 7. sí
9. sí 11. sí 13. sí
15. no; el polinomio cero $\notin H$
17. no; la función $f(x) \equiv 0 \notin V$
19. sí
21. (a) Si A_1 $A_2 \in H_1$, entonces $(A_1 + A_2)_{11} = (A_1)_{11} + (A_2)_{11} = 0 + 0 = 0$, y $(\alpha A_1)_{11} = \alpha(A_1)_{11} = \alpha 0 = 0$ de modo que H_1 es un subespacio. Si $A_1, A_2 \in H_2$,

entonces $A_1 = \begin{pmatrix} -b_1 & a_1 \\ a_1 & b_1 \end{pmatrix}$,

$A_2 = \begin{pmatrix} -b_2 & a_2 \\ a_2 & b_2 \end{pmatrix}$,

$A_1 + A_2$

$= \begin{pmatrix} -(b_1+b_2) & (a_1+a_2) \\ (a_1+a_2) & (b_1+b_2) \end{pmatrix}$

$= \begin{pmatrix} -c & d \\ d & c \end{pmatrix} \in H_2$

y de esta manera H_2 es también un subespacio.

(b) $H = H_1 \cap H_2 = \left\{ A \in M_{22}: A = \begin{pmatrix} 0 & a \\ a & 0 \end{pmatrix} \text{ para algún escalar } a \right\}$. Si A_1, $A_2 \in H$, entonces

$A_1 = \begin{pmatrix} 0 & a_1 \\ a_1 & 0 \end{pmatrix}$, $A_2 = \begin{pmatrix} 0 & a_2 \\ a_2 & 0 \end{pmatrix}$, $A_1 + A_2 = \begin{pmatrix} 0 & a_1+a_2 \\ a_1+a_2 & 0 \end{pmatrix} = \begin{pmatrix} 0 & b \\ b & 0 \end{pmatrix} \in H$ y $\alpha A_1 = \begin{pmatrix} 0 & \alpha a_1 \\ \alpha a_1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & c \\ c & 0 \end{pmatrix} \in H$.

23. Si $\mathbf{x}_1, \mathbf{x}_2 \in H$, entonces $A(\mathbf{x}_1 + \mathbf{x}_2) = A\mathbf{x}_1 + A\mathbf{x}_2 = \mathbf{0} + \mathbf{0} = \mathbf{0}$ de manera que $\mathbf{x}_1 + \mathbf{x}_2 \in H$. Así mismo, $A(\alpha\mathbf{x}_1) = \alpha A\mathbf{x}_1 = \alpha\mathbf{0} = \mathbf{0}$ de forma que $\alpha\mathbf{x}_1 \in H$ y H es un subespacio.

25. Sean $\mathbf{u} = (x_1, y_1, z_1, w_1)$ y $\mathbf{v} = (x_2, y_2, z_2, w_2) \in H$. Entonces $\mathbf{u} + \mathbf{v} = (x_1 + x_2, y_1 + y_2, z_1 + z_2, w_1 + w_2)$ y $a(x_1 + x_2) + b(y_1 + y_2) + c(z_1 + z_2) + d(w_1 + w_2) = (ax_1 + by_1 + cz_1 + dw_1) + (ax_2 + by_2 + cz_2 + dw_2) = 0 + 0 = 0$ de modo que $\mathbf{u} + \mathbf{v} \in H$. En forma análoga, $\alpha\mathbf{u} = (\alpha x_1, \alpha y_1, \alpha z_1, \alpha w_1)$ y $a(\alpha x_1) + b(\alpha y_1) + c(\alpha z_1) + d(\alpha w_1) = \alpha(ax_1 + by_1 + cz_1 + dw_1) = \alpha 0 = 0$ por lo que $\alpha\mathbf{u} \in H$. Consecuentemente H es un subespacio.

27. Sean $\mathbf{x}, \mathbf{y} \in H$. Entonces $\mathbf{x} = \mathbf{u}_1 + \mathbf{v}_1$ y $\mathbf{y} = \mathbf{u}_2 + \mathbf{v}_2$ en donde $\mathbf{u}_1$, y $\mathbf{u}_2 \in H_1$ $\mathbf{v}_1, \mathbf{v}_2 \in H_2$. Entonces $\mathbf{x} + \mathbf{y} = (\mathbf{u}_1 + \mathbf{v}_1) + (\mathbf{u}_2 + \mathbf{v}_2) = (\mathbf{u}_1 + \mathbf{u}_2) + (\mathbf{v}_1 + \mathbf{v}_2)$. Como H_1 y H_2 son subespacios, $\mathbf{u}_1 + \mathbf{u}_2 \in H_1$ y $\mathbf{v}_1 + \mathbf{v}_2 \in H_2$ de modo que $\mathbf{x} + \mathbf{y} \in H$. Similarmente, $\alpha\mathbf{x} = \alpha(\mathbf{u}_1 + \mathbf{v}_1) = \alpha\mathbf{u}_1 + \alpha\mathbf{v}_1$. Pero $\alpha\mathbf{u}_1 \in H_1$ y $\alpha\mathbf{v}_1 \in H_2$ así que $\alpha\mathbf{x} \in H$ y H es un subespacio.

29. Sean $\mathbf{v}_1 = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$ y $\mathbf{v}_2 = \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$. $\mathbf{v}_1$ no es un múltiplo de $\mathbf{v}_2$ puesto que los vectores no son colineales.

Sea $A = \begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix}$. Entonces $\det A = x_1y_2 - x_2y_1$. Si $\det A = 0$, entonces $x_1y_2 = x_2y_1$ o bien $x_1/x_2 = y_1/y_2$ (si $x_2 = 0$ o bien $y_2 = 0$, se puede deducir una conclusión semejante). Sea $c = x_1/x_2 = y_1/y_2$. Entonces $x_1 = cx_2$ y $y_1 = cy_2$, de forma que $\mathbf{v}_1 = c\mathbf{v}_2$ lo cual contradice lo establecido anteriormente. En consecuencia $\det A \neq 0$.

Sea $\mathbf{v} = \begin{pmatrix} x \\ y \end{pmatrix}$ cualquier otro vector en $\mathbb{R}^2$. Deseamos hallar escalares a y b tales que $\mathbf{v} = a\mathbf{v}_1 + b\mathbf{v}_2$, o bien

$\begin{pmatrix} x \\ y \end{pmatrix} = a\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + b\begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$

$= \begin{pmatrix} ax_1 + bx_2 \\ ay_1 + by_2 \end{pmatrix}$

es decir

$\begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}$

o bien

$A\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}$.

Dado que $\det A \neq 0$, este sistema tiene la solución única $\begin{pmatrix} a \\ b \end{pmatrix} = A^{-1}\begin{pmatrix} x \\ y \end{pmatrix}$.

Por consiguiente $\mathbf{v} \in H$, lo cual muestra que $\mathbb{R}^2 \subset H$. Pero como $H \subset \mathbb{R}^2$, tenemos que

$H = \mathbb{R}^2$.

Problemas 4.4

1. independiente

3. dependiente;

$-2\begin{pmatrix} 2 \\ -1 \\ 4 \end{pmatrix} + \begin{pmatrix} 4 \\ -2 \\ 8 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$

5. dependiente (por el Teorema 2)

7. independiente

9. independiente

11. independiente

13. independiente

15. independiente

17. dependiente

19. independiente

21. independiente

23. $ad - bc = 0$ **25.** $\alpha = -\frac{13}{2}$

27. El sistema (7) puede expresarse como

$$(*)\quad c_1\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + c_2\begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} + \cdots + c_n\begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Si (7) tiene incluso una solución no trivial, entonces las columnas de A son linealmente dependientes. Si las columnas de A son dependientes, existen

entonces números $c_1, c_2, \ldots, c_n$ no todos diferentes de cero tales que (*) se verifica.

29. Si $\mathbf{0} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k$, entonces $\mathbf{0} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k + 0\mathbf{v}_{k+1} + 0\mathbf{v}_{k+2} + \cdots + 0\mathbf{v}_n$. Como $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ son independientes, tenemos que $c_1 = c_2 = \cdots = c_k = 0$.

31. Si $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 = 0$, entonces $0 = 0 \cdot \mathbf{v}_1 = (c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3) \cdot \mathbf{v}_1 = c_1(\mathbf{v}_1 \cdot \mathbf{v}_1) + c_2(\mathbf{v}_2 \cdot \mathbf{v}_1) + c_3(\mathbf{v}_3 \cdot \mathbf{v}_1) = c_1\mathbf{v}_1 \cdot \mathbf{v}_1 + c_2 0 + c_3 0 = c_1\mathbf{v}_1 \cdot \mathbf{v}_1$. Dado que $\mathbf{v}_1 \neq 0$, $\mathbf{v}_1 \cdot \mathbf{v}_1 \neq 0$ así que debe tenerse que $c_1 = 0$. Una determinación similar muestra que $c_2 = c_3 = 0$.

33. $x_2\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + x_3\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$

35. $\begin{pmatrix} [illegible] \\ -6 \\ 1 \end{pmatrix}$

$\begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_3\begin{pmatrix} 3 \\ 0 \\ 1 \\ 0 \end{pmatrix} + x_4\begin{pmatrix} -5 \\ 0 \\ 0 \\ 1 \end{pmatrix}$

39. Cualquier $\mathbf{u}$ diferente de cero conduciría a un resultado semejante. Por ejemplo, sea $\mathbf{u} = (1, 0, 0)$. Entonces $\mathbf{x} = (0, 1, 0)$ y $\mathbf{y} = (0, 0, 1)$ están en H. $\mathbf{w} = \mathbf{x} \times \mathbf{y} = (1, 0, 0) = \mathbf{u}$. (e) H es el plano ortogonal a $\mathbf{u}$ en tanto que $\mathbf{w}$ es ortogonal a este plano y por tanto debe ser paralelo a $\mathbf{u}$.

41. Sean $p(x) = a_0 + a_1x + a_2x^2$ y $q(x) = b_0 + b_1x + b_2x^2$ dos polinomios en P_2. Sea $r(x) = c_0 + c_1x + c_2x^2$ un tercer polinomio en P_2. Si p y q generan P_2, entonces existen escalares α y β tales que $r = \alpha p + \beta q$; es decir,

$c_0 = \alpha a_0 + \beta b_0$

$c_1 = \alpha a_1 + \beta b_1$

$c_2 = \alpha a_2 + \beta b_2$

Este sistema "sobredeterminado" de 3 ecuaciones en 2 incógnitas tendrá solución si y sólo si la tercera ecuación es una combinación lineal de las dos primeras. Puesto que c_0, c_1 y c_2 son arbitrarias, rara vez éste será el caso. Por ello p y q no pueden generar P_2. Para ver esto de otra manera, supóngase que $\begin{pmatrix} \alpha \\ \beta \end{pmatrix}$ es una solución del sistema. Entonces

$\begin{pmatrix} c_0 \\ c_1 \end{pmatrix} = \begin{pmatrix} a_0 & b_0 \\ a_1 & b_1 \end{pmatrix}\begin{pmatrix} \alpha \\ \beta \end{pmatrix}$ o bien

$\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} a_0 & b_0 \\ a_1 & b_1 \end{pmatrix}^{-1}\begin{pmatrix} c_0 \\ c_1 \end{pmatrix}$. Pero también, $\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}^{-1}\begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$.

En general estas dos expresiones para $\begin{pmatrix} \alpha \\ \beta \end{pmatrix}$ no serán iguales, lo que significa que, por lo general, el sistema no tiene solución.

43. Sea $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ un subconjunto del conjunto linealmente independiente $T = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ con $k < n$. Supóngase que S es dependiente. Entonces existen escalares diferentes de cero con $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k = \mathbf{0}$. Esto indica que T es dependiente, lo cual es una contradicción. Simplemente fórmese una combinación lineal de los vectores en T, usando c_i siempre que $\mathbf{v}_i \in S$ y 0 en otro caso.

45. Exprese las matrices como $A^{(1)}, A^{(2)}, \ldots, A^{(mn+1)}$. Supóngase que $A^{(i)} = (a_{jk}^{(i)})$. Considere el sistema

$$\begin{aligned} a_{11}^{(1)}\alpha_1 + a_{11}^{(2)}\alpha_2 + \cdots &+ a_{11}^{(mn+1)}\alpha_{mn+1} = 0 \\ \vdots \qquad\quad \vdots \qquad\qquad &\qquad\quad \vdots \qquad\qquad \vdots \\ a_{mn}^{(1)}\alpha_1 + a_{mn}^{(2)}\alpha_2 + \cdots &+ a_{mn}^{(mn+1)}\alpha_{mn+1} = 0 \end{aligned}$$

Este es un sistema homogéneo con ecuaciones mn y $mn + 1$ incógnitas. Consecuentemente tiene una solución diferente de cero y existen escalares $\alpha_1, \alpha_2, \ldots, \alpha_{mn+1}$ no todos nulos tales que $\alpha_1A^{(1)} + \alpha_2A^{(2)} + \cdots + \alpha_{mn+1}A^{mn+1} = 0$ (la matriz cero de $m \times n$).

47. Supóngase que $1, x, \ldots, x^k$ son linealmente independientes. Considere que $c_0 + c_1x + \cdots + c_kx^k + c_{k+1}x^{k+1} = 0$. Si $c_{k+1} \neq 0$, entonces $x^{k+1} = -\frac{c_0}{c_{k+1}} - \frac{c_1}{c_{k+1}}x - \cdots - \frac{c_k}{c_{k+1}}x^k$, lo cual es evidentemente imposible. Por ello $c_{k+1} = 0$. Pero entonces $c_0 + c_1x + \cdots + c_kx^k = 0$, lo cual implica que $c_0 = c_1 = \cdots = c_k = 0$ dado que $1, x, x^2, \ldots, x^k$ son linealmente independientes. Por tanto $1, x, x^2, \ldots, x^k, x^{k+1}$ también son linealmente independientes, y esto completa la demostración por inducción (véase el Apéndice 1).

49. Existen escalares $\alpha_1, \alpha_2,$

$\ldots, \alpha_n$ no todos nulos tales que $\alpha_2 \mathbf{v}_2 + \cdots + \alpha_n \mathbf{v}_n = \mathbf{0}$. Sea k el mayor entero para el cual $\alpha_k \neq 0$ (k puede ser igual a n). Entonces la ecuación queda $\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \cdots + \alpha_k \mathbf{v}_k = \mathbf{0}$ de forma que

$$\mathbf{v}_k = -\frac{\alpha_1}{\alpha_k}\mathbf{v}_1 - \frac{\alpha_2}{\alpha_k}\mathbf{v}_2 - \cdots - \frac{\alpha_{k-1}}{\alpha_k}\mathbf{v}_{k-1}.$$

51. Supóngase que f y g son dependientes. Entonces $g = cf$ y $g' = cf'$ para alguna constante c y

$$w(f, g)(x) = \begin{vmatrix} f(x) & g(x) \\ f'(x) & g'(x) \end{vmatrix} = \begin{vmatrix} f(x) & cf(x) \\ f'(x) & cf'(x) \end{vmatrix} = cf(x)f'(x) - cf(x)f'(x) = 0.$$

53. Considere que $c_1(\mathbf{u} + \mathbf{v}) + c_2(\mathbf{u} + \mathbf{w}) + c_3(\mathbf{v} + \mathbf{w}) = \mathbf{0}$. Entonces $(c_1 + c_2)\mathbf{u} + (c_1 + c_3)\mathbf{v} + (c_2 + c_3)\mathbf{w} = \mathbf{0}$. Como $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ son linealmente independientes,

$$\begin{aligned} c_1 + c_2 \quad &= 0 \\ c_1 \quad + c_3 &= 0 \\ c_2 + c_3 &= 0 \end{aligned}$$

El determinante de este sistema homogéneo es

$$\begin{vmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{vmatrix} = -2 \neq 0,$$

así que la única solución es $c_1 = c_2 = c_3 = 0$ y los tres vectores son independientes

55. $$\begin{vmatrix} 1 & 1 & 1 \\ a & b & c \\ a^2 & b^2 & c^2 \end{vmatrix} = (b-a)(c-a)(c-b)$$

según el resultado del Problema 2.2.35.

Problemas 4.5

1. Sí

3. No; por ejemplo $\begin{pmatrix} 1 \\ 2 \end{pmatrix} \notin$ espacio generado de los tres vectores (cada uno es múltiplo de $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$)

5. sí **7.** sí

9. no; por ejemplo, $x \notin$ gen $\{1-x, 3-x^2\}$

11. sí **13.** sí

15. Sea $p_i(x) = a_{0i} + a_{1i}x + \cdots + a_{ni}x^n$ para $i = 1, 2, \ldots, n$ y defínase el vector

$$\mathbf{a}_i = (a_{0i}, a_{1i}, \ldots, a_{ni}).$$

Elija un vector $\mathbf{b} = (b_0, \ldots, b_n)$ tal que $\mathbf{a}_i \cdot \mathbf{b} = 0$ para $i = 1, 2, \ldots, m$. Si $m < n + 1$, este último sistema homogéneo de ecuaciones siempre tiene una solución no trivial $\mathbf{b}$ (por el Teorema 1.7.1). Supóngase que $\mathbf{b}$ se puede escribir como $\mathbf{b} = \alpha_1\mathbf{a}_1 + \alpha_2\mathbf{a}_2 + \cdots + \alpha_m\mathbf{a}_m$. Entonces $\mathbf{b}\cdot\mathbf{b} = \alpha_1\mathbf{a}_1\cdot\mathbf{b} + \alpha_2\mathbf{a}_2\cdot\mathbf{b} + \cdots + \alpha_m\mathbf{a}_m\cdot\mathbf{b} = 0$. Pero $\mathbf{b} \neq 0$. Esta contradicción muestra que $\mathbf{b}$ no puede expresarse como una combinación lineal de las $\mathbf{a}_i$ si $m < n + 1$. Sea $q(x) = b_0 + b_1x + \cdots + b_nx^n$. Entonces $q(x)$ no puede escribirse como una combinación lineal de los elementos p_i y, por consiguiente las p_i no generan espacio. Se concluye que $m \geq n + 1$.

17. Si $p(x) \in P$, entonces $p(x) = a_0 + a_1x + \cdots + a_kx^k$ para algún entero k. En consecuencia, $p(x)$ puede escribirse como una combinación lineal de $1, x, x^2, \ldots, x^k$.

19. $\alpha\mathbf{v}_1 + \beta\mathbf{v}_2 = (\alpha + \beta c)\mathbf{v}_1$. Sea $\mathbf{v}_1 = (x_1, y_1, z_1)$. Entonces cualquier vector $\mathbf{v} = (x, y, z)$ en gen $\{\mathbf{v}_1, \mathbf{v}_2\}$ puede expresarse como $(x, y, z) = ((\alpha + \beta c)x_1, (\alpha + \beta c)y_1, (\alpha + \beta c)z_1)$, o bien

$$x = (\alpha + \beta c)x_1$$
$$y = (\alpha + \beta c)y_1$$
$$z = (\alpha + \beta c)z_1$$

lo cual, de la Sección 35, es la ecuación de una recta que pasa por (0, 0, 0) con dirección (x_1, y_1, z_1).

21. Sea $\mathbf{v} \in V$. Entonces existen escalares $\alpha_1, \alpha_2, \ldots, \alpha_n$ tales que $\mathbf{v} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_n\mathbf{v}_n = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_n\mathbf{v}_n + 0\mathbf{v}_{n+1}$. La última ecuación muestra que $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{n+1}$ gen V.

23. Sean $\mathbf{v}_i = (v_{i1}, v_{i2}, \ldots, v_{in})$ y $\mathbf{u}_i = (u_{i1}, u_{i2}, \ldots, u_{in})$. Asimismo, sean

$$\mathbf{w}_j = \begin{pmatrix} v_{1j} \\ v_{2j} \\ \vdots \\ v_{nj} \end{pmatrix}, \quad \mathbf{z}_j = \begin{pmatrix} u_{1j} \\ u_{2j} \\ \vdots \\ u_{nj} \end{pmatrix}$$

y $A = (a_{ij})$. Entonces expresando los componentes en las desigualdades dadas, resulta

(*) $$v_{ij} = a_{i1}u_{1j} + a_{i2}u_{2j} + \cdots + a_{in}u_{nj}$$

o bien $\qquad$ dado que $\det A \neq 0$

$\mathbf{w}_j = A\mathbf{z}$; por lo que $\mathbf{z}_j = A^{-1}\mathbf{w}_j$.

Sea $A^{-1} = B = (b_{ij})$. Entonces la expresión para $\mathbf{z}_j$ puede escribirse

$$u_{ij} = b_{i1}v_{1j} + b_{i2}v_{2j} + \cdots + b_{in}v_{nj}.$$

Esto es semejante a la expresión (*) con los elementos u y v intercambiados. Consecuentemente,

$$\mathbf{u}_i = \sum_{k=1}^{n} b_{ik}\mathbf{v}_k \text{ para } i = 1, 2, \ldots, n$$

así que $\mathbf{u}_i \in$ gen $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ para $i = 1, 2, \ldots, n$. Puesto que se dio que $\mathbf{v}_i \in$ gen $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$, concluimos que los espacios generados son iguales.

25. $\begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} -1 \\ 3 \\ 4 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}$. (Hay muchas opciones para el tercer vector.)

Problemas 4.6

1. no; no genera espacio

3. no; dependiente

5. no; no genera espacio

7. sí **9.** sí

11. $\left\{\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}\right\}$ **13.** $\left\{\begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix}\right\}$

15. Como dim $\mathbb{R}^2 = 2$, un subespacio propio H debe tener dimensión 1. Sea $\{(x_0, y_0)\}$ una base para H. Si $(x, y) \in H$, entonces $(x, y) = c(x_0, y_0)$ para algún número real c. Esto significa que $x = cx_0$, $y = cy_0$ o bien

$$c = \frac{x}{x_0} = \frac{y}{y_0} \quad \text{y} \quad y = \left(\frac{y_0}{x_0}\right)x,$$

lo cual es la ecuación de una recta que pasa por el origen y con pendiente $\frac{y_0}{x_0}$ si $x_0 \neq 0$. Si $x_0 = 0$, entonces la recta es el eje y.

17. Sea $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{n-1}\}$ una base para H. Sea $\mathbf{v}_n$ otro vector en H. Entonces los vectores $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{n-1}, \mathbf{v}_n$ son linealmente dependientes. Sea A la matriz cuyos renglones son $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{n-1}, \mathbf{v}_n$. Entonces $\det A = 0$, y la ecuación $A\mathbf{a} = \mathbf{0}$ tiene una solución no trivial

$$\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}.$$

Lo anterior significa que $\mathbf{v}_i \cdot \mathbf{a} = 0$ para $i = 1, 2, \ldots, n$. En particular, si $\mathbf{v}_n = (x_1, x_2, \ldots, x_n)$, entonces $\mathbf{v}_n \cdot \mathbf{a} = 0$, o bien $a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0$. Dado que $\mathbf{v}_n$ era un vector arbitrario en H, esto prueba el resultado.

19. $\left\{\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right\}$ **21.** $\left\{\begin{pmatrix} -2 \\ -3 \\ 1 \end{pmatrix}\right\}$

23. $\left\{\begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}\right\}$ **25.** n

27. V tiene una base $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ así que existen escalares tales que

$$\begin{aligned} \mathbf{v}_1 &= a_{11}\mathbf{u}_1 + a_{12}\mathbf{u}_2 + \cdots \\ \mathbf{v}_2 &= a_{21}\mathbf{u}_1 + a_{22}\mathbf{u}_2 + \cdots \\ &\vdots \\ \mathbf{v}_m &= a_{m1}\mathbf{u}_1 + a_{m2}\mathbf{u}_2 + \cdots \\ &\quad + a_{1n}\mathbf{u}_n \\ &\quad + a_{2n}\mathbf{u}_n \\ &\quad \vdots \\ &\quad + a_{mn}\mathbf{u}_n \end{aligned}$$

Sea $\mathbf{a}_i = (a_{i1}, a_{i2}, \ldots, a_{in})$. Los elementos $\mathbf{a}_i$ son m vectores linealmente independientes en $\mathbb{R}^2$ (de otro modo los elementos no serían independientes). Desarrolle los elementos $\mathbf{a}_i$ en una base $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m, \mathbf{a}_{m+1}, \ldots, \mathbf{a}_n$ para $\mathbb{R}^n$ sumando simplemente al conjunto $n - m$ vectores linealmente independientes. Entonces si $\mathbf{a}_k = (a_{k1}, a_{k2}, \ldots, a_{kn})$ para $m < k \leq n$, defina $\mathbf{v}_k = a_{k1}\mathbf{u}_1 + a_{k2}\mathbf{u}_2 + \cdots + a_{kn}\mathbf{u}_n$ para $k = m + 1, m + 2, \ldots, n$. Puesto que

$$\det\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \neq 0,$$

el conjunto $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ forma una base para V ya que consiste en n vectores linealmente independientes en V con dim $V = n$.

29. Si los vectores son independientes, entonces forman una base y dim $V = n$. Si no es así, entonces por el Problema 4.4.49, por lo menos uno de ellos se puede expresar como una combinación lineal de los que lo preceden. Descarte este vector. Siga de este modo hasta que queden m vectores linealmente independientes. Éstos aún deben generar V por la forma en que fueron elegidos. Por lo tanto, dim $V = m < n$. En cualquier caso dim $V \leq n$.

31. (a) Vea el Problema 4.3.27. (b) Sea $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ una base para H, y $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\}$ una base para K. Claramente $B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n, \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\}$ genera $H + K$. Suponga que $\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_n\mathbf{v}_n + \beta_1\mathbf{u}_1 + \beta_2\mathbf{u}_2 + \cdots + \beta_m\mathbf{u}_m = \mathbf{0}$ en donde no todos los coeficientes son cero. Sean $\mathbf{h} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_n\mathbf{v}_n$ y $\mathbf{k} = \beta_1\mathbf{u}_1 + \beta_2\mathbf{u}_2 + \cdots + \beta_m\mathbf{u}_m$.

Entonces por independencia lineal, ni $\mathbf{h}$ ni $\mathbf{k}$ son el vector cero. Asimismo, $\mathbf{h} \in H$ y $\mathbf{k} \in K$. Pero entonces $\mathbf{h} + \mathbf{k} = \mathbf{0}$ o bien $\mathbf{h} = -\mathbf{k} \in K$. Consecuentemente $\mathbf{0} \neq \mathbf{h} \in H \cap K$, lo cual

contradice el hecho de que $H \cap K = \{\mathbf{0}\}$. Por consiguiente, son nulos todos los elementos α y β, lo que implica que los vectores en B son linealmente independientes. Por ello B es una base para $H + K$ y $\dim(H + K) = \dim H + \dim K$.

(i) Si $\dim \text{gen}\{\mathbf{v}_1, \mathbf{v}_2\} = 1$, entonces elíjase una base $\{\mathbf{v}\}$ para $\text{gen}\{\mathbf{v}_1, \mathbf{v}_2\}$. Entonces $\mathbf{v}_1 = \alpha\mathbf{v}$ y $\mathbf{v}_2 = \beta\mathbf{v}$. Si $\mathbf{v}_1 = 0$, entonces $\mathbf{v}_1$ es un único punto que está en $\mathbf{v}_2$. Si $\mathbf{v}_1 \neq \mathbf{0}$, entonces $\alpha \neq 0$ por lo que $\mathbf{v}_2 = \beta\mathbf{v} = \frac{\beta}{\alpha}(\alpha\mathbf{v}) = \frac{\alpha}{\beta}\mathbf{v}_1$. En cualquier caso los vectores son colineales.
(ii) Si los vectores son colineales, entonces $\mathbf{v}_2 = c\,\mathbf{v}_1$ para algún escalar c de modo que $\mathbf{v}_1$ es una base para $\text{gen}\{\mathbf{v}_1, \mathbf{v}_2\}$ y $\dim \text{gen}\{\mathbf{v}_1, \mathbf{v}_2\} = 1$.

35. Si no son linealmente independientes, entonces, como en la respuesta al Problema 29, $\dim V < n$. Como $\dim V = n$, los vectores deben ser independientes y, por consiguiente, constituyen una base para V.

37. $B_1 = \left\{\begin{pmatrix}1\\0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\\1\end{pmatrix}, \begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\\0\end{pmatrix}\right\}$

$B_2 = \left\{\begin{pmatrix}1\\0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\\1\end{pmatrix}, \begin{pmatrix}0\\0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\\1\end{pmatrix}\right\}$

Existen muchísimas otras opciones.

Problemas 4.7

1. $\rho = 2,\ \nu = 0$ **3.** $\rho = 1,\ \nu = 2$
5. $\rho = 2,\ \nu = 1$ **7.** $\rho = 2,\ \nu = 2$
9. $\rho = 2,\ \nu = 0$ **11.** $\rho = 2,\ \nu = 2$
13. $\rho = 3,\ \nu = 1$ **15.** $\rho = 2,\ \nu = 1$

17. base imagen $= \left\{\begin{pmatrix}1\\3\\5\end{pmatrix}, \begin{pmatrix}-1\\1\\-1\end{pmatrix}\right\}$; éstas son las primeras dos columnas de A.

base espacio nulo $= \left\{\begin{pmatrix}-3\\1\\2\end{pmatrix}\right\}$

19. base imagen $= \left\{\begin{pmatrix}1\\0\\1\end{pmatrix}, \begin{pmatrix}-1\\1\\0\end{pmatrix}, \begin{pmatrix}3\\3\\5\end{pmatrix}\right\}$; éstas son las tres primeras columnas (linealmente independientes) de A.

base espacio nulo $= \left\{\begin{pmatrix}-6\\-4\\1\\0\end{pmatrix}\right\}$

21. base imagen $= \left\{\begin{pmatrix}1\\-2\\2\\3\end{pmatrix}\right\}$

base espacio nulo $= \left\{\begin{pmatrix}1\\1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\2\\1\\0\end{pmatrix}, \begin{pmatrix}0\\3\\0\\1\end{pmatrix}\right\}$

23. $\left\{\begin{pmatrix}1\\4\\-2\end{pmatrix}, \begin{pmatrix}0\\1\\-\frac{6}{7}\end{pmatrix}\right\}$

25. $\{(1, 0, 0, \frac{1}{2}), (0, 1, -1, \frac{3}{2}), (0, 0, 0, 1)\}$

27. no **29.** sí

31. Si $\mathbf{c}_i$ denota la i-ésima columna de D, entonces

$$\mathbf{c}_i = d_i\begin{pmatrix}0\\0\\\vdots\\1\\0\\\vdots\\0\end{pmatrix} \leftarrow i\text{-ésima posición}$$

i-ésima posición. En consecuencia, los elementos $\mathbf{c}_i$ son linealmente independientes cuando $d_i \neq 0$, y el número de columnas linealmente independientes es el rango.

33. $\rho(A^t) =$ dimensión del espacio columna de A^t = dimensión del espacio renglón de A = dimensión del espacio de columnas de A (por el Teorema 3) $= \rho(A)$.

35. (i) Sea H = imagen de A y sea $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ una base para H. Ya que B es invertible, $\ker B = \{\mathbf{0}\}$, lo que significa que $\{B\mathbf{v}_1, B\mathbf{v}_2, \ldots, B\mathbf{v}_k\}$ es un conjunto linealmente independiente en $\mathbb{R}^m$ y en consecuencia es una base para Imag BA. Por consiguiente $\rho(BA) = k = \rho(A)$.
(ii) Dado que C es invertible, Imag $C = \mathbb{R}^n$. Sea $\mathbf{h} \in H$; entonces existe un $\mathbf{x} \in \mathbb{R}^n$ tal que $A\mathbf{x} = \mathbf{h}$. Ya que Imag $C = \mathbb{R}^n$, existe un $\mathbf{y} \in \mathbb{R}^n$ tal que $C\mathbf{y} = \mathbf{x}$. Consecuentemente $AC\mathbf{y} = \mathbf{h}$. Por ello $H \subset$ Imag AC. Si $\mathbf{v} \in$ Imag AC, existe un $\mathbf{u}$ en $\mathbb{R}^n$ tal que $AC\mathbf{u} = \mathbf{v}$. Pero entonces $\mathbf{v} = A(C\mathbf{u})$ de manera que $\mathbf{v} \in$ Imag $A = H$. Por lo tanto Imag $AC \subset H$ de modo que Imag $AC = H$ y $\rho(A) = \rho(AC)$.

37. Puesto que $\rho(A) = 5$, los cinco renglones de A son linealmente independientes. En consecuencia, los cinco renglones de $(A, \mathbf{b})$ son linealmente independientes, y $\rho(A, \mathbf{b}) = 5$.

39. Por el Problema 35, $\rho(A) = \rho(AD) = \rho(C(AD)) = \rho(B)$.

41. (i) Si existe un $\mathbf{x} \neq \mathbf{0}$ tal que $A\mathbf{x} = \mathbf{0}$, entonces $A(\alpha\mathbf{x}) = \alpha A\mathbf{x} = \mathbf{0}$ para todo $\alpha \in \mathbb{R}$ de modo que $\nu(A) = \dim \ker A \geq 1$, y $\rho(A) = n - \nu(A) \leq n - 1 < n$.
(ii) Si $\rho(A) < n$, entonces $\nu(A) = n - \rho(A) > 0$ así que existe un $\mathbf{x} \neq \mathbf{0}$ tal que $A\mathbf{x} = \mathbf{0}$.

Problemas 4.8

1. 1 **3.** 2 **5.** 2 **7.** 3 **9.** 3

11. Sea $|M|$ un subdeterminante de orden $k + 1$. Desarrolle $|M|$ en su primer renglón. Al hacer esto obtenemos $k + 1$ determinantes de orden k. Como cada determinante es cero por hipótesis, $|M| = 0$.

Problemas 4.9

1. $\frac{x+y}{2}\begin{pmatrix}1\\1\end{pmatrix} + \frac{x-y}{2}\begin{pmatrix}1\\-1\end{pmatrix} = \begin{pmatrix}x\\y\end{pmatrix}$

3. $\frac{4x+3y}{41}\begin{pmatrix}5\\7\end{pmatrix} + \frac{7x-5y}{41}\begin{pmatrix}3\\-4\end{pmatrix} = \begin{pmatrix}x\\y\end{pmatrix}$

5. $\frac{dx-by}{ad-bc}\begin{pmatrix}a\\b\end{pmatrix} + \frac{-cx+ay}{ad-bc}\begin{pmatrix}b\\d\end{pmatrix} = \begin{pmatrix}x\\y\end{pmatrix}$

7. $(x-y)\begin{pmatrix}1\\0\\0\end{pmatrix} + (y-z)\begin{pmatrix}1\\1\\0\end{pmatrix} + z\begin{pmatrix}1\\1\\1\end{pmatrix} = \begin{pmatrix}x\\y\\z\end{pmatrix}$

9. $\frac{6x-11y+10z}{31}\begin{pmatrix}2\\1\\3\end{pmatrix} + \frac{2x+17y-7z}{31}\begin{pmatrix}-1\\4\\5\end{pmatrix} + \frac{7x+13y-9z}{31}\begin{pmatrix}3\\-2\\-4\end{pmatrix} = \begin{pmatrix}x\\y\\z\end{pmatrix}$

11. $a_0 + a_1x + a_2x^2 = (a_0 + a_1 + a_2)1 + a_1(x-1) + a_2(x^2-1)$

13. $a_0 + a_1x + a_2x^2 = \frac{(a_0+a_1+a_2)}{2}(x+1) + \frac{(a_1-a_0-a_2)}{2}(x-1) + a_2(x^2-1)$

15. $2(x^3+x^2) - 5(x^2+x) + 10(x+1) - 16(1)$

17. $(\mathbf{x})_{B_2} = \begin{pmatrix}-\frac{1}{3}\\0\end{pmatrix}$

19. $(\mathbf{x})_{B_2} = \begin{pmatrix}\frac{86}{33}\\-\frac{20}{11}\\\frac{7}{11}\end{pmatrix}$

21. independiente
23. dependiente
25. independiente
27. independiente

29. Si fueran linealmente independientes, generarían P_n. Pero $1 \in P_n$ y $1 \notin \text{gen}\{p_1, p_2, \ldots, p_{n+1}\}$ puesto que es 0 el término constante en cada polinomio.

31. Si fueran linealmente independientes, generarían M_{mn}. Pero la matriz $A = (a_{ij})$ en donde $a_{11} = 1$ y $a_{ij} = 0$ de otro modo no está en el espacio generado de $A_1, A_2, \ldots, A_{mn+1}$ ya que una combinación lineal de matrices con un 0 en la posición 1, 1 también tiene un cero en la posición 1, 1.

33. $\begin{pmatrix}\cos\theta & -\text{sen}\,\theta\\ \text{sen}\,\theta & \cos\theta\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}\cos\theta\\ \text{sen}\,\theta\end{pmatrix}$;

$\begin{pmatrix}\cos\theta & -\text{sen}\,\theta\\ \text{sen}\,\theta & \cos\theta\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}-\text{sen}\,\theta\\ \cos\theta\end{pmatrix}$.

A^{-1} se obtiene girando un ángulo de $-\theta$. Por tanto

$A^{-1} = \begin{pmatrix}\cos(-\theta) & -\text{sen}(-\theta)\\ \text{sen}(-\theta) & \cos(-\theta)\end{pmatrix} = \begin{pmatrix}\cos\theta & \text{sen}\,\theta\\ -\text{sen}\,\theta & \cos\theta\end{pmatrix}$.

De manera alternativa, dado que $A = \begin{pmatrix}\cos\theta & -\text{sen}\,\theta\\ \text{sen}\,\theta & \cos\theta\end{pmatrix}$,

$A^{-1} = \begin{pmatrix}\cos\theta & \text{sen}\,\theta\\ -\text{sen}\,\theta & \cos\theta\end{pmatrix}$.

35. $A^{-1} = \begin{pmatrix}\cos(\pi/4) & \text{sen}(\pi/4)\\ -\text{sen}(\pi/4) & \cos(\pi/4)\end{pmatrix} = \begin{pmatrix}1/\sqrt{2} & 1/\sqrt{2}\\ -1/\sqrt{2} & 1/\sqrt{2}\end{pmatrix}$

(del Problema 33) así que

$A^{-1}\begin{pmatrix}2\\-7\end{pmatrix} = \begin{pmatrix}-5/\sqrt{2}\\-9/\sqrt{2}\end{pmatrix}$

37. Como C es invertible, las columnas de C son linealmente independientes. Esto es, $\{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n\}$ son n vectores linealmente independientes en V, los cuales son, en consecuencia, una base para V puesto que $\dim V = n$.

39. Si $CA = I$, entonces $(\mathbf{x})_{B_1} = I(\mathbf{x})_{B_1} = CA(\mathbf{x})_{B_1}$. Recíprocamente, suponga que $(\mathbf{x})_{B_1} = CA(\mathbf{x})_{B_1}$. Sea $B_1 = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$. Entonces

$$(\mathbf{v}_1)_{B_1} = \begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix} = CA\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}.$$

Sea

$$CA = \begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1n}\\ r_{21} & r_{22} & \cdots & r_{2n}\\ \vdots & \vdots & & \vdots\\ r_{n1} & r_{n2} & \cdots & r_{nn}\end{pmatrix}.$$

Entonces $CA\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}$ = la primera columna de

$$CA = \begin{pmatrix}r_{11}\\r_{22}\\\vdots\\r_{n1}\end{pmatrix} = \begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}.$$

Análogamente, la segunda columna de $CA = \begin{pmatrix}0\\1\\0\\\vdots\\0\end{pmatrix}$ ya que

$$(\mathbf{v}_2)_{B_1} = \begin{pmatrix}0\\1\\0\\\vdots\\0\end{pmatrix}.$$ Continuando de esta forma, se ve que $CA = I$.

Problemas 4.10

1. $\begin{pmatrix}1/\sqrt{2}\\1/\sqrt{2}\end{pmatrix}, \begin{pmatrix}-1/\sqrt{2}\\1/\sqrt{2}\end{pmatrix}$

3. (i) Si $a = b = 0$, $\{(1, 0), (0, 1)\}$
 (ii) Si $a = 0$, $b \neq 0$ $\{(1, 0)\}$
 (iii) Si $a \neq 0$, $b = 0$ $\{(0, 1)\}$
 (iv) Si $a \neq 0$, $b \neq 0\{(b/\sqrt{a^2+b^2}, -a/\sqrt{a^2+b^2})\}$

5. $\{(1/\sqrt{5}, 0, 2/\sqrt{5}), (2/\sqrt{30}, 5/\sqrt{30}, -1/\sqrt{30})\}$

7. $\{(2/\sqrt{29}, 3/\sqrt{29}, 4/\sqrt{29})\}$

9. $\{(1/\sqrt{5}, 0, 0, 2/\sqrt{5}), (2/\sqrt{30}, 5/\sqrt{30}, 0, -1/\sqrt{30}), (-2/\sqrt{10}, 1/\sqrt{10}, 2/\sqrt{10}, 1/\sqrt{10})\}$

11. $\{(a/\sqrt{a^2+b^2+c^2}, b/\sqrt{a^2+b^2+c^2}, c/\sqrt{a^2+b^2+c^2})\}$

13. $\{(-7/\sqrt{66}, -1/\sqrt{66}, 4/\sqrt{66})\}$

15. $Q^t = \begin{pmatrix}\frac{2}{3} & \frac{1}{3} & -\frac{2}{3}\\ \frac{1}{3} & \frac{2}{3} & \frac{2}{3}\\ \frac{2}{3} & -\frac{2}{3} & \frac{1}{3}\end{pmatrix}$ y $Q^tQ = I = QQ^t$

17. $PQ = \dfrac{1}{3\sqrt{2}} \times \begin{pmatrix}1-\sqrt{8} & -1-\sqrt{8}\\ 1+\sqrt{8} & 1-\sqrt{8}\end{pmatrix}$,
 $(PQ)^t = \dfrac{1}{3\sqrt{2}} \times \begin{pmatrix}1-\sqrt{8} & 1+\sqrt{8}\\ -1-\sqrt{8} & 1-\sqrt{8}\end{pmatrix}$,
 $(PQ)(PQ)^t = \dfrac{1}{18}\begin{pmatrix}18 & 0\\ 0 & 18\end{pmatrix} = I$

19. $I = Q^{-1}Q = Q^tQ = QQ = Q^2$. Pero $\det Q^2 = (\det Q)^2 = \det I = 1$, de modo que $\det Q = \pm 1$.

21. Si $\mathbf{v}_i = 0$, entonces $0\mathbf{v}_1 + 0\mathbf{v}_2 + \cdots + 0\mathbf{v}_{i-1} + \mathbf{v}_i + 0\mathbf{v}_{i+1} + \cdots + 0\mathbf{v}_n = \mathbf{0}$, lo que implica que los elementos $\mathbf{v}_i$ son linealmente dependientes. Por ello $\mathbf{v}_i \neq 0$ para $i = 1, 2, \ldots, n$.

23. (a) $\mathbf{0}$
 (b) $\dfrac{1}{\sqrt{a^2+b^2}}\begin{pmatrix}a\\b\end{pmatrix}$
 (c) $\mathbf{v} = \begin{pmatrix}a\\b\end{pmatrix} + \begin{pmatrix}0\\0\end{pmatrix}$

25. (a) $\dfrac{1}{49}\begin{pmatrix}-186\\75\\118\end{pmatrix}$ (b) $\dfrac{1}{7}\begin{pmatrix}3\\-2\\6\end{pmatrix}$
 (c) $\mathbf{v} = \dfrac{1}{49}\begin{pmatrix}-186\\75\\118\end{pmatrix} + \dfrac{13}{49}\begin{pmatrix}3\\-2\\6\end{pmatrix}$

27. (a) $\dfrac{1}{5}\begin{pmatrix}1\\-3\\4\\17\end{pmatrix}$
 (b) $\dfrac{1}{\sqrt{15}}\begin{pmatrix}2\\-1\\3\\-1\end{pmatrix}$
 (c) $\dfrac{1}{5}\begin{pmatrix}1\\-3\\4\\17\end{pmatrix} + \dfrac{2}{5}\begin{pmatrix}2\\-1\\3\\-1\end{pmatrix}$

29. $|\mathbf{u}_1 - \mathbf{u}_2|^2 = (\mathbf{u}_1 - \mathbf{u}_2) \cdot (\mathbf{u}_1 - \mathbf{u}_2) = \mathbf{u}_1 \cdot \mathbf{u}_1 - \mathbf{u}_2 \cdot \mathbf{u}_1 - \mathbf{u}_1 \cdot \mathbf{u}_2 + \mathbf{u}_2 \cdot \mathbf{u}_2 = 1 - 0 - 0 + 1 = 2$ en virtud de que $\mathbf{u}_1$, $\mathbf{u}_2$ son ortonormales.

31. $a^2 + b^2 = 1$

33. $0 \leq \left(\dfrac{\mathbf{u}}{|\mathbf{u}|} - \dfrac{\mathbf{v}}{|\mathbf{v}|}\right) \cdot \left(\dfrac{\mathbf{u}}{|\mathbf{u}|} - \dfrac{\mathbf{v}}{|\mathbf{v}|}\right)$

$$= \frac{\mathbf{u}\cdot\mathbf{u}}{|\mathbf{u}|^2} - \frac{2\mathbf{u}\cdot\mathbf{v}}{|\mathbf{u}||\mathbf{v}|} + \frac{\mathbf{v}\cdot\mathbf{v}}{|\mathbf{v}|^2}$$

$$= \frac{|\mathbf{u}|^2}{|\mathbf{u}|^2} - \frac{2\mathbf{u}\cdot\mathbf{v}}{|\mathbf{u}||\mathbf{v}|} + \frac{|\mathbf{v}|^2}{|\mathbf{v}|^2}$$

$$= 2 - \frac{2\mathbf{u}\cdot\mathbf{v}}{|\mathbf{u}||\mathbf{v}|} \geq 0$$

así que $\dfrac{\mathbf{u}\cdot\mathbf{v}}{|\mathbf{u}||\mathbf{v}|} \leq 1$ y $\mathbf{u}\cdot\mathbf{v} \leq |\mathbf{u}||\mathbf{v}|$. (1) De modo semejante,

$$0 \leq \left(\frac{\mathbf{u}}{|\mathbf{u}|} + \frac{\mathbf{v}}{|\mathbf{v}|}\right) \cdot \left(\frac{\mathbf{u}}{|\mathbf{u}|} + \frac{\mathbf{v}}{|\mathbf{v}|}\right)$$

$= 2 + \dfrac{2\mathbf{u}\cdot\mathbf{v}}{|\mathbf{u}||\mathbf{v}|} \geq 0$ de forma que $1 \geq -\dfrac{\mathbf{u}\cdot\mathbf{v}}{|\mathbf{u}||\mathbf{v}|}$ o bien

$|\mathbf{u}||\mathbf{v}| \geq -\mathbf{u}\cdot\mathbf{v}$. (2)
Combinando (1) y (2) se tiene que $|\mathbf{u}\cdot\mathbf{v}| \leq |\mathbf{u}||\mathbf{v}|$.

35. $|\mathbf{u} + \mathbf{v}|^2 = (|\mathbf{u}| + |\mathbf{v}|)^2$. Esto significa que $(\mathbf{u} + \mathbf{v}) \cdot (\mathbf{u} + \mathbf{v}) = |\mathbf{u}|^2 + 2\mathbf{u}\cdot\mathbf{v} + |\mathbf{v}|^2 = (|\mathbf{u}| + |\mathbf{v}|)^2 = |\mathbf{u}|^2 + 2|\mathbf{u}||\mathbf{v}| + |\mathbf{v}|^2$. Por consiguiente $\mathbf{u}\cdot\mathbf{v} = |\mathbf{u}||\mathbf{v}|$ lo cual, del Problema 34, puede ocurrir sólo si $\mathbf{u} = \lambda\mathbf{v}$; es decir, $\mathbf{u}$ y $\mathbf{v}$ son linealmente dependientes.

37. Se prueba esto por inducción matemática. Si $k = 2$, éste es el resultado del Problema 35. Suponemos que se verifica para $k = n$ y lo demostramos para $k = n + 1$. Supóngase que $|\mathbf{x}_1 + \mathbf{x}_2 + \cdots + \mathbf{x}_n + \mathbf{x}_{n+1}| = |\mathbf{x}_1| + |\mathbf{x}_2| + \cdots + |\mathbf{x}_n| + |\mathbf{x}_{n+1}|$. (*) Esto implica que $|\mathbf{x}_1 + \mathbf{x}_2 + \cdots + \mathbf{x}_n| = |\mathbf{x}_1| + |\mathbf{x}_2| + \cdots + |\mathbf{x}_n|^{(\surd)}$, debido a que si esto no se verifica, entonces, por la desigualdad del triángulo,

$$|\mathbf{x}_1 + \mathbf{x}_2 + \cdots + \mathbf{x}_n| < |\mathbf{x}_1| + |\mathbf{x}_2| + \cdots + |\mathbf{x}_n|.$$

Pero entonces

$$|\mathbf{x}_1 + \mathbf{x}_2 + \cdots + \mathbf{x}_n + \mathbf{x}_{n+1}| \leq |\mathbf{x}_1 + \mathbf{x}_2 + \cdots + \mathbf{x}_n| + |\mathbf{x}_{n+1}| < |\mathbf{x}_1| + |\mathbf{x}_2| + \cdots + |\mathbf{x}_n| + |\mathbf{x}_{n+1}|,$$

lo cual contradice (*). Consecuentemente, por la hipótesis de la inducción, dim gen $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\} = 1$. Sea $\mathbf{u} = \mathbf{x}_1 + \mathbf{x}_2 + \cdots + \mathbf{x}_n$. Por (*) y ($\surd$) $|\mathbf{u} + \mathbf{x}_{n+1}| = |\mathbf{u}| + |\mathbf{x}_{n+1}|$ de modo que por el Problema 35, $\mathbf{x}_{n+1} = \lambda\mathbf{u}$ para algún número λ. Es decir, $\mathbf{x}_{n+1} \in$ gen $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ por lo que dim gen $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n, \mathbf{x}_{n+1}\} = 1$ también. Por tanto, el resultado es cierto para $k = n + 1$, y la demostración queda completa.

39. $(H^\perp)^\perp = \{\mathbf{v} \in \mathbb{R}^n; \mathbf{v}\cdot\mathbf{k} = 0$ para todo $\mathbf{k} \in H^\perp\}$. Sea $\mathbf{x} \in H$; entonces $\mathbf{x}\cdot\mathbf{k} = 0$ para todo $\mathbf{k} \in H^\perp$ por lo cual $\mathbf{x} \in (H^\perp)^\perp$, lo que muestra que $H \subseteq (H^\perp)^\perp$. Recíprocamente, si $\mathbf{v} \in (H^\perp)^\perp$, entonces $\mathbf{v}\cdot\mathbf{k} = 0$ para todo $\mathbf{k} \in H^\perp$. Pero $\mathbf{v} = \mathbf{h}' + \mathbf{k}'$ en donde $\mathbf{h}' \in H$ y $\mathbf{k}' \in H^\perp$. Entonces $0 = \mathbf{v}\cdot\mathbf{k} = \mathbf{h}'\cdot\mathbf{k} + \mathbf{k}'\cdot\mathbf{k} = 0 + \mathbf{k}'\cdot\mathbf{k}$. Por ello $\mathbf{k}'\cdot\mathbf{k} = 0$ para todo $\mathbf{k} \in H$, lo que significa, en particular, que $\mathbf{k}'\cdot\mathbf{k}' = 0$. En consecuencia, $\mathbf{k}' = 0$ y $\mathbf{v} = \mathbf{h}' \in H$. De este modo, $(H^\perp)^\perp \subset H$ y, junto con $H \subset (H^\perp)^\perp$, demuestra que $(H^\perp)^\perp = H$.

41. Sea $\mathbf{k} \in H_2^\perp$. Entonces $\mathbf{k}\cdot\mathbf{h} = 0$ para todo $\mathbf{h} \in H_2$. Como $H_1 \subset H_2$, esto demuestra que $\mathbf{k}\cdot\mathbf{h} = 0$ para todo $\mathbf{h} \in H_1$. Esto es, $\mathbf{k} \in H_1^\perp$. Por lo tanto $H_2^\perp \subset H_1^\perp$.

Problemas 4.11

1. (i) $(A, A) = a_{11}^2 + a_{22}^2 + \cdots + a_{nn}^2 \geq 0$.
 (ii) $(A, A) = 0$ implica que $a_{ii}^2 = 0$ para $i = 1, 2, \ldots, n$ por lo que $A = 0$. Si $A = 0$, entonces $(A, A) = 0$.
 (iii) $(A, B + C) = a_{11}(b_{11} + c_{11}) + \cdots + a_{nn}(b_{nn} + c_{nn}) = a_{11}b_{11} + a_{11}c_{11} + \cdots + a_{nn}b_{nn} + a_{nn}c_{nn} = (a_{11}b_{11} + \cdots + a_{nn}b_{nn}) + (a_{11}c_{11} + \cdots + a_{nn}c_{nn}) = (A, B) + (A, C)$
 (iv) De igual forma, $(A + B, C) = (A, C) + (B, C)$
 (v) $(A, B) = (B, A) = \overline{(B, A)}$ puesto que todos los componentes son reales y $a_{ii}b_{ii} = b_{ii}a_{ii}$
 (vi) $(\alpha A, B) = (\alpha a_{11})b_{11} + \cdots + (\alpha a_{nn})b_{nn} = \alpha[a_{11}b_{11} + \cdots + a_{nn}b_{nn}] = \alpha(A, B)$
 (vii) $(A, \alpha B) = \overline{(\alpha B, A)} = (\alpha B, A) = \alpha(B, A) = \alpha(A, B) = \alpha(A, B)$

3. Sea E_i la matriz de $n \times n$ con un 1 en la posición i, i, y 0 en todas las demás posiciones. Es fácil ver que $\{E_1, E_2, \ldots, E_n\}$ es una base ortonormal para D_n.

5. $\{(1/\sqrt{2}, i/\sqrt{2}), (i/\sqrt{2}, 1/\sqrt{2})\}$

7. $\{\frac{1}{2}, \sqrt{3/2}x, \sqrt{5/8}\,(3x^2 - 1)\}$

9. Primeramente nótese que si $A =$ y $B^t = (b_{ji})$, entonces

$$(AB^t)_{ij} = \sum_{k=1}^{n} a_{ik}b_{jk}$$

de manera que

$$tr(AB^t) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}b_{ij}.$$

 (i) $(A, A) = tr(AA^t) = \sum_{i=1}^{n}\left(\sum_{j=1}^{n} a_{ij}^2\right) \geq 0$
 (ii) $(A, A) = 0$ implica que $a_{ij}^2 = 0$ para toda i y toda j por lo que $A = 0$. Si $A = 0$, entonces $A^t = 0$ y $AA^t = 0$ y así $tr(AA^t) = 0$.
 (iii) $(A, B + C) = tr[A(B + C)^t] + tr[A(B^t + C^t)] = tr(AB^t + AC^t) = tr(AB^t) + tr(AC^t) = (A, B) + (A, C)$
 (iv) De forma análoga $(A + B, C) = (A, C) + (B, C)$
 (v) $(A, B) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}b_{ij} = tr(BA^t) = (B, A)$
 (vi) $(\alpha A, B) = tr(\alpha AB^t) = \alpha tr(AB^t) = \alpha(A, B)$
 (vii) $(A, \alpha B) = (\alpha B, A) = \alpha(B, A) = \alpha(A, B)$

11. $\begin{pmatrix}1 & 0\\0 & 0\end{pmatrix}, \begin{pmatrix}0 & 1\\0 & 0\end{pmatrix}, \begin{pmatrix}0 & 0\\1 & 0\end{pmatrix}, \begin{pmatrix}0 & 0\\0 & 1\end{pmatrix}$

13. (a) (i) $(p, p) = p(a)^2 + p(b)^2 + p(c)^2 \geq 0$
 (ii) $(p, p) = 0$ implica que $p(a) = p(b) = p(c) = 0$. Pero una cuadrática puede tener a lo sumo dos raíces. Por ello $p(x) = 0$ para

toda x. Recíprocamente si $p \equiv 0$, entonces $p(a) = p(b) = p(c) = 0$ por lo cual $(p, p) = 0$.

(iii) $(p, q + r)$

$$= p(a)(q(a) + r(a)) + p(b)(q(b) + r(b)) + p(c)(q(c) + r(c))$$

$$= [p(a)q(a) + p(b)q(b) + p(c)q(c)] + [p(a)r(a) + p(b)r(b) + p(c)r(c)]$$

$$= (p, q) + (p, r)$$

(iv) Similarmente, $(p + q, r) = (p, r) + (q, r)$

(v) $(p, q) = p(a)q(a) + (p(b)q(b) + p(c)q(c)$

$$= q(a)p(a) + q(b)p(b) + q(c)p(c)$$

$$= (q, p)$$

(vi) $(\alpha p, q) = [\alpha p(a)]q(a) + [\alpha p(b)]q(b) + [\alpha p(c)]q(c)$

$$= \alpha[p(a)q(a) + p(b)q(b) + p(c)q(c)]$$

$$= \alpha(p, q)$$

(vii) $(p, \alpha q) = (\alpha p, q) = \alpha(p, q) = \alpha(p, q)$

(b) No, dado que se infringe **(ii)**. Por ejemplo, sean $a = 1$, $b = -1$ y $p(x) = (x - 1)(x + 1) = x^2 - 1 \neq 0$. Entonces $p(a) = p(b) = 0$ de manera que $(p, p) = 0$ aun cuando $p \neq 0$. De hecho, para cualquier polinomio q, se tiene que $(p, q) = 0$.

15. $\sqrt{31}$

17. $0 \leq \left(\left(\frac{\mathbf{u}}{|\mathbf{u}|} - \frac{\mathbf{v}}{|\mathbf{v}|}\right), \left(\frac{\mathbf{u}}{|\mathbf{u}|} - \frac{\mathbf{v}}{|\mathbf{v}|}\right)\right)$

$$= \frac{(\mathbf{u}, \mathbf{u})}{|\mathbf{u}|^2} - \frac{(\mathbf{u}, \mathbf{v})}{|\mathbf{u}||\mathbf{v}|} - \frac{(\mathbf{v}, \mathbf{u})}{|\mathbf{u}||\mathbf{v}|} + \frac{(\mathbf{v}, \mathbf{v})}{|\mathbf{v}|^2}$$

$$= \frac{|\mathbf{u}|^2}{|\mathbf{u}|^2} - \left[\frac{(\mathbf{u}, \mathbf{v}) + \overline{(\mathbf{u}, \mathbf{v})}}{|\mathbf{u}||\mathbf{v}|}\right] + \frac{|\mathbf{v}|^2}{|\mathbf{v}|^2}.$$

Ahora bien, si $z = a + bi$, entonces $z + \bar{z} = (a + bi) + (a - bi) = 2a = 2 \operatorname{Re} z$ (y $z - \bar{z} = 2bi = 2iImz$). Por consiguiente $(\mathbf{u}, \mathbf{v}) + \overline{(\mathbf{u}, \mathbf{v})} = 2 \operatorname{Re}(\mathbf{u}, \mathbf{v})$ y así $2 - \frac{2 \operatorname{Re}(\mathbf{u}, \mathbf{v})}{|\mathbf{u}||\mathbf{v}|} \geq 0$ o bien $\frac{\operatorname{Re}(\mathbf{u}, \mathbf{v})}{|\mathbf{u}||\mathbf{v}|} \leq 1$. Sea λ un número real. Entonces $0 \leq ((\lambda\mathbf{u} + (\mathbf{u}, \mathbf{v})\mathbf{v}), (\lambda\mathbf{u} + (\mathbf{u}, \mathbf{v})\mathbf{v})) = \lambda^2|\mathbf{u}|^2 + |(\mathbf{u}, \mathbf{v})|^2|\mathbf{v}|^2 + \lambda\overline{(\mathbf{u}, \mathbf{v})}(\mathbf{u}, \mathbf{v}) + \bar{\lambda}(\mathbf{u}, \mathbf{v})(\mathbf{v}, \mathbf{u}) =$ (pues λ es real) $\lambda^2|\mathbf{u}|^2 + 2\lambda|(\mathbf{u}, \mathbf{v})|^2 + |(\mathbf{u}, \mathbf{v})|^2|\mathbf{v}|^2$. Lo último es una ecuación cuadrática en λ. Si $a\lambda^2 + b\lambda + c \geq 0$, entonces la ecuación $a\lambda^2 + b\lambda + c = 0$ puede tener como máximo una raíz real y, por lo tanto, $b^2 - 4ac \leq 0$. En consecuencia

$$4((|\mathbf{u}, \mathbf{v})|^2)^2 - 4|\mathbf{u}|^2 \times |(\mathbf{u}, \mathbf{v})|^2|\mathbf{v}|^2 \leq 0$$

o bien $|(\mathbf{u}, \mathbf{v})|^2 \leq |\mathbf{u}|^2|\mathbf{v}|^2$

y $\quad |(\mathbf{u}, \mathbf{v})| \leq |\mathbf{u}||\mathbf{v}|$.

19. $H^{\perp} =$ gen $\{(-15x^2 + 16x - 3), (20x^3 - 30x^2 + 12x - 1)\}$

21. $1 + 2x + 3x^2 - x^3$

$$= \frac{30x^2 + 52x + 19}{20} + \frac{(-20x^3 + 30x^2 - 12x + 1)}{20}$$

23. $\frac{2}{\pi} + \sqrt{3}\left(\frac{2}{\pi} - \frac{8}{\pi^2}\right)\sqrt{3}(2x - 1) + \sqrt{5}\left(\frac{2}{\pi} + \frac{24}{\pi^2} - \frac{96}{\pi^3}\right) \times \sqrt{5}(6x^2 - 6x + 1)$

$$\approx -0.8346x^2 - 0.2091x + 1.1585$$

25. $A^* = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{2} - \frac{i}{2} & \frac{1}{2} + \frac{i}{2} \end{pmatrix}$.

Verifique que $A^*A = I$.

Problemas 4.12

1. Sea L un conjunto linealmente independiente en V. Sea S la colección de todos los subconjuntos de V linealmente independientes, ordenados parcialmente, por inclusión tal que todo conjunto en S contenga a L. La demostración entonces se desarrolla como en la del Teorema 2.

3. El resultado es cierto para $n = 2$. Supóngase que se verifica para $n = k$. Considere los $k + 1$ conjuntos $A_1, A_2, \ldots, A_k, A_{k+1}$ en una cadena. Los primeros k conjuntos forman una cadena y, por la hipótesis de inducción, uno de ellos contiene a los otros $k - 1$ conjuntos. Llámese a este conjunto A_i. Entonces $A_i \subseteq A_{k+1}$ o bien $A_{k+1} \subseteq A_i$. En uno u otro casos hemos encontrado un conjunto que contiene los otros k conjuntos y el resultado es válido para $n = k + 1$. Esto completa la demostración por inducción.

Ejercicios de repaso • Capítulo 4.

1. sí; dimensión 2; base $\{(1, 0, 1), (0, 1, 2)\}$

3. sí; dimensión 3; base $\{(1, 0, 0, -1), (0, 1, 0, -1), (0, 0, 1, -1)\}$

5. sí; dimensión $[n(n+1)]/2$; base $\{(E_{ij} : j \geq 1\}$ en donde E_{ij} es la matriz con un 1 en la posición i, j y 0 en las demás.

7. no; por ejemplo $(x^5 - 2x) + (-x^5 + x^2) = x^2 - 2x$, lo cual no es un polinomio de grado 5, así que el conjunto no es cerrado ante la adición.

9. no; por ejemplo,

$$\begin{pmatrix} 1 & 1 \\ 0 & 2 \\ 3 & 1 \end{pmatrix} + \begin{pmatrix} 2 & 1 \\ -1 & 2 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 3 & 2 \\ -1 & 4 \\ 4 & 1 \end{pmatrix},$$ lo cual no satisface $a_{12} = 1$.

11. independiente

13. dependiente

15. independiente

17. independiente

19. independiente

21. dimensión 2; base $\{(2, 0, 1), (0, 4, 3)\}$

23. dimensión 3; base $\{(1, 0, 3, 0), (0, 1, -1, 0), (0, 0, 1, 1)\}$

25. dimensión 4; base $\{D_1, D_2, D_3, D_4\}$ en donde D_i es la matriz con un 1 en la posición i, i y 0 en las demás

27. Imag $A =$ gen $\left\{\begin{pmatrix} 1 \\ -2 \end{pmatrix}\right\}$;

ker $A =$ gen $\left\{\begin{pmatrix} 2 \\ 1 \end{pmatrix}\right\}$;

$\rho(A) = \nu(A) = 1$

29. Imag $A = \mathbb{R}^3$;
ker $A = \{\mathbf{0}\}$;
$\rho(A) = 3$, $\nu(A) = 0$

31. Imag A

$$= \text{gen}\left\{\begin{pmatrix} 2 \\ -1 \\ 4 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \\ 6 \end{pmatrix}\right\};$$

ker $A = \{\mathbf{0}\}$;
$\rho(T) = 2$, $\nu(T) = 0$

33. $\frac{3}{4}\begin{pmatrix} 1 \\ 2 \end{pmatrix} - \frac{5}{4}\begin{pmatrix} -1 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$

35. $1(1 + x^2) + 0(1 + x) + 3(1) = 4 + x^2$

37. 2

39. $\left\{\frac{1}{\sqrt{13}}\begin{pmatrix} 2 \\ 3 \end{pmatrix}, \frac{1}{\sqrt{13}}\begin{pmatrix} -3 \\ 2 \end{pmatrix}\right\}$

41. $\begin{pmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{pmatrix}$

43. (a) $\begin{pmatrix} \frac{4}{3} \\ -\frac{1}{3} \\ \frac{5}{3} \end{pmatrix}$ (b) $\begin{pmatrix} -1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{pmatrix}$

(c) $\begin{pmatrix} \frac{4}{3} \\ -\frac{1}{3} \\ \frac{5}{3} \end{pmatrix} + \begin{pmatrix} -\frac{7}{3} \\ \frac{7}{3} \\ \frac{7}{3} \end{pmatrix}$

45. (a) $\begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$

(b) $\left\{\begin{pmatrix} 1/\sqrt{2} \\ 0 \\ -1/\sqrt{2} \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1/\sqrt{2} \\ 0 \\ -1/\sqrt{2} \end{pmatrix}\right\}$

(c) $\begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix} + \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$

47. $\frac{e^2}{2} + 3(x - 1) + (\frac{15}{4}e^2 - \frac{105}{4})(x^2 - 2x + \frac{2}{3}) \approx 1.6672 + 0.0821x + 1.459x^2$

Capítulo 5

Problemas 5.1

1. lineal **3.** lineal

5. no lineal, puesto que

$$T\left(\alpha\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = T\begin{pmatrix} \alpha x \\ \alpha y \\ \alpha z \end{pmatrix} = \begin{pmatrix} 1 \\ \alpha z \end{pmatrix}$$

mientras que

$$T\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \alpha\begin{pmatrix} 1 \\ z \end{pmatrix} = \begin{pmatrix} \alpha \\ \alpha z \end{pmatrix}$$

7. lineal

9. no lineal, ya que

$$T\left(\alpha\begin{pmatrix} x \\ y \end{pmatrix}\right) = T\begin{pmatrix} \alpha x \\ \alpha y \end{pmatrix} = (\alpha x)(\alpha y) = \alpha^2 xy$$

en tanto que $\alpha T\begin{pmatrix} x \\ y \end{pmatrix} = \alpha xy$

11. lineal

13. no lineal, ya que

$$T\left(\alpha\begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix}\right) = \alpha^2 T\begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} \neq \alpha T\begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix}$$

si $\alpha \neq 1$ o bien 0

15. no lineal, pues

$$T(A + B) = (A + B)^t(A + B) = (A^t + B^t)(A + B) = A^tA + A^tB + B^tA + B^tB.$$

Pero

$$T(A) + T(B) = A^tA + B^tB \neq T(A + B)$$

a menos que $A^tB + B^tA = 0$.

17. no lineal, pues $T(\alpha D) = (\alpha D)^2 = \alpha^2 D^2 \neq \alpha T(D) = \alpha D^2$ a menos que $\alpha = 1$ o bien 0.

19. lineal **21.** lineal

23. no lineal, pues $T(f + g) = (f + g)^2 \neq f^2 + g^2 = T(f) + T(g)$

25. lineal **27.** lineal

29. no lineal, pues

$$T(\alpha A) = \det(\alpha A) = \alpha^n \det A \neq \alpha \det A = \alpha T(A)$$

a menos que $\alpha = 0$ o bien 1. [det $\alpha A = \alpha^n$ det A por

el Problema 3.2.28.] También, en general, $\det(A+B) \neq \det A + \det B$.

31. (a) $\begin{pmatrix} -14 \\ 4 \\ 26 \end{pmatrix}$ (b) $\begin{pmatrix} -31 \\ -6 \\ 26 \end{pmatrix}$

33. Gira un vector en sentido contrario al del reloj alrededor del eje z un ángulo θ en un plano paralelo al plano xy.

35. Supóngase que $\alpha < 0$. Entonces $T[(\alpha - \alpha)\mathbf{x}] = T(0\mathbf{x}) = 0T\mathbf{x} = \mathbf{0}$. Por tanto $T[(\alpha - \alpha)\mathbf{x}] = T(0\mathbf{x}) = 0T\mathbf{x} = 0$ y $T(\alpha\mathbf{x}) + T(-\alpha\mathbf{x}) = T((\alpha - \alpha)\mathbf{x}) = \mathbf{0}$. Pero $-\alpha > 0$ de modo que $T(-\alpha\mathbf{x}) = -\alpha T\mathbf{x}$. En consecuencia $T(\alpha\mathbf{x}) - \alpha T\mathbf{x} = 0$, o bien $T(\alpha\mathbf{x}) = \alpha T\mathbf{x}$ para $\alpha < 0$ también.

37. $T(\mathbf{x} - \mathbf{y}) = T\mathbf{x} + T(-\mathbf{y}) = T\mathbf{x} + T[(-1)\mathbf{y}] = T\mathbf{x} + (-1)T\mathbf{y} = T\mathbf{x} - T\mathbf{y}$.

39. $T(\mathbf{v}_1 + \mathbf{v}_2) = (\mathbf{v}_1 + \mathbf{v}_2, \mathbf{u}_0) = (\mathbf{v}_1, \mathbf{u}_0) + (\mathbf{v}_2, \mathbf{u}_0) = T\mathbf{v}_1 + T\mathbf{v}_2$; $T(\alpha\mathbf{v}) = (\alpha\mathbf{v}, \mathbf{u}_0) = \alpha(\mathbf{v}, \mathbf{u}_0) = \alpha T\mathbf{v}$.

41.
$$\begin{aligned} T(\mathbf{v}_1 + \mathbf{v}_2) &= (\mathbf{v}_1 + \mathbf{v}_2, \mathbf{u}_1)\mathbf{u}_1 = (\mathbf{v}_1 + \mathbf{v}_2, \mathbf{u}_2)\mathbf{u}_2 + \cdots + (\mathbf{v}_1 + \mathbf{v}_2, \mathbf{u}_n)\mathbf{u}_n \\ &= (\mathbf{v}_1, \mathbf{u}_1)\mathbf{u}_1 + (\mathbf{v}_2, \mathbf{u}_1)\mathbf{u}_1 + (\mathbf{v}_1, \mathbf{u}_2)\mathbf{u}_2 + (\mathbf{v}_2, \mathbf{u}_2)\mathbf{u}_2 + \cdots + (\mathbf{v}_1, \mathbf{u}_n)\mathbf{u}_n + (\mathbf{u}_2, \mathbf{u}_n)\mathbf{u}_n \\ &= (\mathbf{v}_1, \mathbf{u}_1)\mathbf{u}_1 + (\mathbf{v}_1, \mathbf{u}_2)\mathbf{u}_2 + \cdots + (\mathbf{v}_1, \mathbf{u}_n)\mathbf{u}_n + (\mathbf{v}_2, \mathbf{u}_1)\mathbf{u}_1 + (\mathbf{v}_2, \mathbf{u}_2)\mathbf{u}_2 + \cdots + (\mathbf{v}_2, \mathbf{u}_n)\mathbf{u}_n \\ &= T\mathbf{v}_1 + T\mathbf{v}_2; \end{aligned}$$
$$\begin{aligned} T(\alpha\mathbf{v}) &= (\alpha\mathbf{v}, \mathbf{u}_1)\mathbf{u}_1 + (\alpha\mathbf{v}, \mathbf{u}_2)\mathbf{u}_2 + \cdots + (\alpha\mathbf{v}, \mathbf{u}_n)\mathbf{u}_n \\ &= \alpha[(\mathbf{v}, \mathbf{u}_1)\mathbf{u}_1 + (\mathbf{v}, \mathbf{u}_2)\mathbf{u}_2 + \cdots + (\mathbf{v}, \mathbf{u}_n)\mathbf{u}_n] \\ &= \alpha T\mathbf{v}. \end{aligned}$$

Problemas 5.2

1. kernel (o núcleo) $= \{(0, y) : y \in \mathbb{R}\}$, es decir, el eje y; imag $= \{(x, 0) : x \in \mathbb{R}\}$, es decir, el eje x; $\rho(T) = \nu(T) = 1$

3. kernel $= \{(x, -x) : x \in \mathbb{R}\}$— esto es la recta $x + y = 0$; imag $= \mathbb{R}$; $\rho(T) = \nu(T) = 1$.

5. kernel $= \left\{\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}\right\}$; imagen $= M_{22}$; $\rho(T) = 4$, $\nu(T) = 0$

7. kernel $= \{A : A^t = -A\} = \{A : A$ es antisimétrica- imagen A es simétrica$\}$; imag $= \{A : A$ es simétrica$\}$; $\rho(T) = (n^2+n)/2$; $\nu(T) = (n^2-n)/2$

9. kernel (o núcleo) $= \{f \in C[0, 1] : f(\frac{1}{2}) = 0\}$; imagen $= \mathbb{R}$; $\rho(T) = 1$; el kernel o núcleo es un espacio de dimensión infinita, de modo que $\nu(T) = \infty$. Por ejemplo, las funciones linealmente independientes $x - \frac{1}{2}$, $(x - \frac{1}{2})^2$, $(x - \frac{1}{2})^3$, $(x - \frac{1}{2})^4, \ldots, (x - \frac{1}{2})^n, \ldots$ satisfacen todas $f(\frac{1}{2}) = 0$.

11. Si $\mathbf{v} \in V$, entonces $\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n$ de modo que $T\mathbf{v} = T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n) = c_1T\mathbf{v}_1 + c_2T\mathbf{v}_2 + \cdots + c_nT\mathbf{v}_n) = c_1\mathbf{0} + c_2\mathbf{0} + \cdots + c_n\mathbf{0} = \mathbf{0}$. Así $T\mathbf{v} = \mathbf{0}$ para todo $\mathbf{v} \in V$ y es, por lo tanto, la transformación cero.

13. La imagen de T es un subespacio de $\mathbb{R}^3$ y, por el Ejemplo 4.6.9, los subespacios de $\mathbb{R}^3$ son $\{0\}$, $\mathbb{R}^3$, y rectas y planos que pasan por el origen.

15. $T\mathbf{x} = A\mathbf{x}$ en donde $A = \begin{pmatrix} 0 & a \\ b & c \end{pmatrix}$, a, b, c son reales

17. $T\mathbf{x} = A\mathbf{x}$ en donde $A = \begin{pmatrix} 2 & -1 & 1 \\ 2 & -1 & 1 \\ 2 & -1 & 1 \end{pmatrix}$

19. (i) Si $A \in \ker T$, entonces $A - A^t = 0$, o bien $A = A^t$.
(ii) Si $A \in$ imagen de T, entonces hay una matriz B tal que $B - B^t = A$. Asimismo $A^t = (B - B^t)^t = B^t - (B^t)^t = B^t - B = -A$ de modo que A es antisimétrica.

21. Sea $T_{ij}(\mathbf{u}_i) = \mathbf{w}_j$ y $T_{ij}(\mathbf{u}_k) = \mathbf{0}$ si $k \neq i$. Estas forman una base para $L(V, W)$, así $\dim L(V, W) = \mathrm{nm}$.

23. Falso. Sean S y T: $\mathbb{R}^2 \to \mathbb{R}^2$ dados por $S(\mathbf{x}) = A\mathbf{x}$ y $T(\mathbf{x}) = B\mathbf{x}$, en donde $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ y $B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$. Entonces $ST(\mathbf{x}) = AB\mathbf{x} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}\mathbf{x} = \mathbf{0}$. Sin embargo, $TS(\mathbf{x})$ no es una transformación cero porque $BA = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \neq$ la matriz cero.

Problemas 5.3

1. $\begin{pmatrix} 1 & -2 \\ -1 & 1 \end{pmatrix}$; $\ker T = \{\mathbf{0}\}$; imag $T = \mathbb{R}^2$; $\nu(T) = 0$, $\rho(T) = 2$

3. $\begin{pmatrix} 1 & -1 & 1 \\ -2 & 2 & -2 \end{pmatrix}$; imag $T =$ gen $\left\{\begin{pmatrix} 1 \\ -2 \end{pmatrix}\right\}$; $\ker T =$ gen $\left\{\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}\right\}$; $\rho(T) = 1$, $\nu(T) = 2$

5. $\begin{pmatrix} 1 & -1 & 2 \\ 3 & 1 & 4 \\ 5 & -1 & 8 \end{pmatrix}$; imag $T =$ gen $\left\{\begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix}\right\}$; $\ker T =$ gen $\left\{\begin{pmatrix} -3 \\ 1 \\ 2 \end{pmatrix}\right\}$; $\rho(T) = 2$, $\nu(T) = 1$

7. $\begin{pmatrix} 1 & -1 & 2 & 3 \\ 0 & 1 & 4 & 3 \\ 1 & 0 & 6 & 6 \end{pmatrix}$; imag $T =$ gen $\left\{\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}\right\}$; $\ker T =$ gen $\left\{\begin{pmatrix} -6 \\ -4 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -6 \\ -3 \\ 0 \\ 1 \end{pmatrix}\right\}$; $\rho(T) = 2$, $\nu(T) = 2$

9. $\begin{pmatrix} \frac{5}{4} & -\frac{13}{4} \\ \frac{5}{4} & \frac{3}{4} \end{pmatrix}$; imag $T = \mathbb{R}^2$; $\ker T = \{\mathbf{0}\}$; $\rho(T) = 2$, $\nu(T) = 0$

11. $\begin{pmatrix} 3 & \frac{7}{5} & \frac{16}{5} \\ 0 & \frac{4}{5} & \frac{2}{5} \end{pmatrix}$; imag $T = \mathbb{R}^2$; $\ker T =$ gen $\left\{\begin{pmatrix} 5 \\ 3 \\ -6 \end{pmatrix}_{B_1}\right\}$; $\rho(T) = 2$, $\nu(T) = 1$

13. $\begin{pmatrix} 0 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$; imag $T =$ gen $\{1 - x, x^3\}$; $\ker T =$ gen $\{x^2\}$; $\rho(T) = 2$, $\nu(T) = 1$

15. $(0, 0, 1, 0)$; imag $T = \mathbb{R}$; $\ker T =$ gen $\{1, x, x^3\}$; $\rho(T) = 1$, $\nu(T) = 3$

17. $\begin{pmatrix} 1 & -1 & 2 & 3 \\ 0 & 1 & 4 & 3 \\ 1 & 0 & 6 & 5 \end{pmatrix}$; imag $T =$ gen $\{1 + x^2, -1 + x, 2 + 4x + 6x^2\} = P_2$; $\ker T =$ gen $\{x^2 - 4x - 6\}$; $\rho(T) = 3$, $\nu(T) = 1$

19. $\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$; imag $T = M_{22}$; $\ker T = \left\{\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}\right\}$; $\rho(T) = 4$, $\nu(T) = 0$

21. $\begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 4 \end{pmatrix}$; imag $D = P_3$; $\ker D = \mathbb{R}$; $\rho(D) = 4$, $\nu(D) = 1$

23. $\begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 2 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & 0 & \cdots & n \\ 0 & 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$; imag $D = P_{n-1}$; $\ker D = \mathbb{R}$; $\rho(D) = n$, $\nu(D) = 1$

25. $\begin{pmatrix} 2 & 0 & 2 & 0 & 0 \\ 0 & 3 & 0 & 6 & 0 \\ 0 & 0 & 4 & 0 & 12 \\ 0 & 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 0 & 6 \end{pmatrix}$; imag $T = P_4$; $\ker T = \{\mathbf{0}\}$; $\rho(T) = 5$, $\nu(T) = 0$

27. $A_T = \mathrm{diag}(b_0, b_1, b_2, \ldots, b_n)$ en donde $b_j = \sum_{i=1}^{j+1} \frac{j!}{(j+1-i)!}$; imag $T = P_n$; $\ker T = \{\mathbf{0}\}$; $\rho(T) = n + 1$, $\nu(T) = 0$

29. $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$; imag $T = P_2$; $\ker T = \{\mathbf{0}\}$, $\rho(T) = 3$, $\nu(T) = 0$

31. Por ejemplo, en M_{34},

$$A_T = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

En general, $A_T = (a_{ij})$ en donde

$$a_{ij} = \begin{cases} 1, & \text{si } i = km + l, \\ & \text{y } j = (l-1)n + k + 1 \\ & \text{para } k = 1, 2, \ldots, n-1 \\ & \text{y } l = 1, 2, \ldots, m \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

33. $\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}$; imag D $=$ gen $\{\text{sen}\, x, \cos x\}$; $\ker D = \mathbb{R}$; $\rho(D) = 2$, $\nu(D) = 1$

35. $\begin{pmatrix} \frac{1}{2} & -i/2 \\ i/2 & \frac{1}{2} \end{pmatrix}$

37. Sean B_1 y B_2 bases para V y W, respectivamente. Tenemos $(T\mathbf{v})_{B_2} = A_T(\mathbf{v})_{B_1}$ para todo $\mathbf{v} \in V$. Entonces $\mathbf{v} \in \ker T$ si y sólo si $T\mathbf{v} = \mathbf{0}$ si y sólo si $A_T(\mathbf{v})_{B_1} = (\mathbf{0})_{B_2}$ si y sólo si $(\mathbf{v})_{B_1} \in \ker A_T$. Por tanto, nulidad de $T = N_{A_T}$ y así $\nu(T) = \nu(A_T)$. Si $\mathbf{w} \in$ imag T, entonces $T\mathbf{v} = \mathbf{w}$ para algún $\mathbf{v} \in V$ de manera que $A_T(\mathbf{v})_{B_1} = (T\mathbf{v})_{B_2} = (\mathbf{w})_{B_2}$. Esto significa que $(\mathbf{w})_{B_2} \in R_{A_T}$. Así $R_{A_T} =$ imag T y así $\rho(T) = \rho(A_T)$. Como $\nu(A_T) + \rho(A_T) = n$ por el Teorema 4.7.5, vimos que $\nu(T) + \rho(T) = n$ también.

Problemas 5.4

1. Como $(\alpha A)^t = \alpha A^t$ y $(A + B)^t = A^t + B^t$, T es lineal. $A^t = 0$ si y sólo si $A = 0$, de modo que $\ker T = \{0\}$ y T es 1-1. Para cualquier matriz A, $(A^t)^t = A$, por lo que T es sobre.

3. (i) Si T es un isomorfismo, entonces $T\mathbf{x} = A_T\mathbf{x} = \mathbf{0}$ si y sólo si $\mathbf{x} = \mathbf{0}$. Por tanto, por el Teorema Resumen, $\det A_T \neq 0$. (ii) Si $\det A_T \neq 0$, entonces $A_T\mathbf{x} = \mathbf{0}$ tiene solamente solución trivial. Así, T es 1-1 y, puesto que V y W tienen dimensión infinita, T es también sobre.

5. $m = [n(n+1)]/2 = \dim\{A : A \text{ es } n \times n \text{ y simétrica}\}$.

7. Defínase $T: P_4 \to W$ por $Tp = xp$. $Tp = 0$ implica $p(x) = 0$; esto es, p es el polinomio cero. Consecuentemente, T es 1-1 y, puesto que $\dim W = 5$, T es también sobre.

9. $mn = pq$

11. La demostración del Teorema 6 prueba la afirmación con la consideración de que los escalares $c_1, c_2, \ldots, c_n$ son números complejos.

13. $T(A_1 + A_2) = (A_1 + A_2)B = A_1B + A_2B = TA_1 + TA_2$; $T(\alpha A) = (\alpha A)B = \alpha(AB) = \alpha TA$. Por tanto T es lineal. Supóngase que $TA = 0$. Entonces $AB = 0$. Como B es invertible, podemos multiplicar por la izquierda por B^{-1} para obtener $A = ABB^{-1} = 0B^{-1} = 0$ o bien $A = 0$. Así, T es $1 - 1$, y como $\dim M_{nn} = n^2 < \infty$, T es un isomorfismo.

15. Elíjase $\mathbf{h} \in H$. Entonces $\text{Proy}_H\, \mathbf{h} = \mathbf{h}$ de modo que T es sobre. Si $H = V$, entonces T es también 1 a 1.

17. Puesto que T es un isomorfismo, $\ker T = \ker A = \{0\}$ de manera que, por el Teorema Resumen A es invertible. Si $\mathbf{x} = T^{-1}\mathbf{y}$, entonces $T\mathbf{x} = A\mathbf{x} = \mathbf{y}$ de modo que $\mathbf{x} = A^{-1}\mathbf{y}$ pues A^{-1} existe. Por consiguiente $T^{-1}\mathbf{y} = A^{-1}\mathbf{y}$ para todo $\mathbf{y} \in \mathbb{R}^n$.

19. Para $z = a + ib \in \mathbb{C}$, defínase $Tz = (a, b) \in \mathbb{R}^2$. Entonces $T(z_1 + z_2) = T((a_1 + a_2) + i(b_1 + b_2)) = (a_1 + a_2, b_1 + b_2) = (a_1, b_1) + (a_2, b_2) = Tz_1 + Tz_2$. Si $\alpha \in \mathbb{R}$, entonces $T(\alpha z) = T(\alpha(a + ib)) = T(\alpha a + i\alpha b) = (\alpha a, \alpha b) = \alpha(a, b) = \alpha Tz$. Por tanto T es lineal. Finalmente, si $T(z) = (0, 0)$, entonces claramente $z = a + ib = 0 + i0 = 0$. Así, T es 1 a 1 y puesto que $\dim \mathbb{C}$ (sobre los reales) $= \dim \mathbb{R}^2 = 2$, T es un isomorfismo.

21. Del Teorema 1.11.1′, el mapeo $T\mathbf{x} = A_T\mathbf{x}$ es sobre si y sólo si A_T tiene inversa derecha.

Problemas 5.5

1. $T\mathbf{x} \cdot T\mathbf{y}$

$$= \begin{pmatrix} x_1 \,\text{sen}\,\theta + x_2 \cos\theta \\ x_1 \cos\theta - x_2 \,\text{sen}\,\theta \\ x_3 \end{pmatrix} \cdot \begin{pmatrix} y_1 \,\text{sen}\,\theta + y_2 \cos\theta \\ y_2 \cos\theta - y_2 \,\text{sen}\,\theta \\ y_3 \end{pmatrix}$$

$$= x_1y_1(\text{sen}^2\theta + \cos^2\theta) + x_2y_2(\text{sen}^2\theta + \cos^2\theta) + x_3y_3$$

$$= x_1y_1 + x_2y_2 + x_3y_3 = \mathbf{x} \cdot \mathbf{y}$$

(se suprimen todos los otros términos del producto escalar).

3. Utilizando el Teorema 1, $T\mathbf{x} \cdot T\mathbf{y} = (AB\mathbf{x}) \cdot (AB\mathbf{y}) = \mathbf{x} \cdot (AB)^t(AB\mathbf{y}) = \mathbf{x} \cdot (B^tA^t)(AB)\mathbf{y} = \mathbf{x} \cdot (B^{-1}A^{-1}AB)\mathbf{y} = \mathbf{x} \cdot \mathbf{y}$.

5. $|\mathbf{x} + \mathbf{y}|^2 = (\mathbf{x} + \mathbf{y}) \cdot (\mathbf{x} + \mathbf{y}) = \mathbf{x} \cdot \mathbf{x} + 2\mathbf{x} \cdot \mathbf{y} + \mathbf{y} \cdot \mathbf{y} = |\mathbf{x}|^2 + 2\mathbf{x} \cdot \mathbf{y} + |\mathbf{y}|^2$ de modo que $\mathbf{x} \cdot \mathbf{y} = \frac{1}{2}(|\mathbf{x} + \mathbf{y}|^2 - |\mathbf{x}|^2 - |\mathbf{y}|^2)$; $T\mathbf{x} \cdot T\mathbf{y} = \frac{1}{2}(|T\mathbf{x} + T\mathbf{y}|^2 - |T\mathbf{x}|^2 - |T\mathbf{y}|^2) = \frac{1}{2}(|T(\mathbf{x} + \mathbf{y})|^2 - |T\mathbf{x}|^2 - |T\mathbf{y}|^2) = \frac{1}{2}(|\mathbf{x} + \mathbf{y}|^2 - |\mathbf{x}|^2 - |\mathbf{y}|^2)$ (puesto que $|T\mathbf{x}| = |\mathbf{x}|$) $= \mathbf{x} \cdot \mathbf{y}$

7. $T\mathbf{x} = \alpha\mathbf{x}$ en donde α es un escalar y $\alpha \neq 0$ o bien 1.

9. $T\mathbf{x} \cdot T\mathbf{y} = \mathbf{x} \cdot \mathbf{y} = A\mathbf{x} \cdot A\mathbf{y}$ y $A^t = A^{-1}$ de modo que $A = (A^{-1})^t$. Entonces $\mathbf{x} \cdot \mathbf{y} = \mathbf{x} \cdot (I\mathbf{y}) = \mathbf{x} \cdot (A^{-1})^tA^{-1}\mathbf{y} = A^{-1}\mathbf{x} \cdot A^{-1}\mathbf{y} = S\mathbf{x} \cdot S\mathbf{y}$ de manera que $S\mathbf{x} = A^{-1}\mathbf{y}$ es una isometría.

11. $T(a_0 + a_1x + a_2x^2 + a_3x^3) = (a_0/\sqrt{2} - (5/2\sqrt{2})a_2, \sqrt{\tfrac{3}{2}}a_1 - (3\sqrt{7}/2\sqrt{2})a_3, (3\sqrt{5}/2\sqrt{2})a_2, (5\sqrt{7}/2\sqrt{2})a_3)$

13. $T\begin{pmatrix} a & b \\ c & d \end{pmatrix} = (a/\sqrt{2} - (5/2\sqrt{2})c, \sqrt{\tfrac{3}{2}}b - (3\sqrt{7}/2\sqrt{2})d, (3\sqrt{5}/2\sqrt{2})c, (5\sqrt{7}/2\sqrt{2})d)$

15. $A^* = \begin{pmatrix} 1-i & 3 \\ -4-2i & 6+3i \end{pmatrix}$

17. Si A es hermitiana, entonces $A^* = A$. En particular, las componentes diagonales de A no se mueven cuando se toma la transpuesta, de modo que $\overline{a_{ii}} = a_{ii}$, lo cual significa que a_{ii} es real.

19. Sea $A^* = B = (b_{ij})$ y sea $\mathbf{c}_i$ la columna i de A. Entonces $AB = I = (\delta_{ij})$ en donde $\delta_{ij} = \begin{cases} 1, \text{si } i = j \\ 0, \text{si } i \neq j \end{cases}$. Pero $\delta_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj} = \sum_{k=1}^{n} a_{ik}\overline{a_{jk}} = \mathbf{c}_i \cdot \mathbf{c}_j = \delta_{ij}$.

21. Como la componente i de $A\mathbf{x}$ es $\sum_{j=1}^{n} a_{ij}x_j$, se tiene que $(A\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_j\overline{y_i}$. Análogamente, si $A^* = B = (b_{ij})$, $(x, A^*y) =$

$$\sum_{i=1}^{n}\sum_{j=1}^{n} x_j\overline{b_{ji}y_i} = \sum_{j=1}^{n}\sum_{i=1}^{n} x_j\overline{b_{ji}y_i} = \sum_{j=1}^{n}\sum_{i=1}^{n} x_ja_{ij}\overline{y_i} = \sum_{i=1}^{n}\sum_{j=1}^{n} \times a_{ij}x_j\overline{y_i} = (A\mathbf{x}, \mathbf{y}).$$

Ejercicios de repaso • Capítulo 5

1. lineal

3. no lineal, pues $T(\alpha(x, y)) = T(\alpha x, \alpha y) = \alpha x/\alpha y = x/y = T(x, y) \neq \alpha T(x, y)$ a menos que $\alpha = 1$.

5. no lineal, ya que $T(p_1 + p_2) = 1 + p_1 + p_2$, pero $Tp_1 + Tp_2 = (1 + p_1) + (1 + p_2) = 2 + p_1 + p_2$.

7. $\begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}$; imag T $=$ gen $\left\{\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right\}$; $\ker T =$ gen $\left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right\}$; $\rho(T) = \nu(T) = 1$

9. $\begin{pmatrix} 1 & 0 & -2 & 0 \\ 0 & 2 & 0 & 3 \end{pmatrix}$; imag $T = \mathbb{R}^2$; $\ker T$ $=$ gen $\left\{\begin{pmatrix} 2 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -3 \\ 0 \\ 2 \end{pmatrix}\right\}$; $\rho(A) = \nu(A) = 2$

11. $\begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}$; imag $T = M_{22}$; $\ker T = \{\mathbf{0}\}$; $\rho(T) = 4$, $\nu(T) = 0$

13. $T(a_0 + a_1x + a_2x^2) = \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix}$

Capítulo 6

Problemas 6.1

1. $-4, 3$; $E_{-4} =$ gen $\left\{\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right\}$; $E_3 =$ gen $\left\{\begin{pmatrix} 2 \\ -5 \end{pmatrix}\right\}$

3. $i, -i$; $E_i =$ gen $\left\{\begin{pmatrix} 2+i \\ 5 \end{pmatrix}\right\}$; $E_{-i} =$ gen $\left\{\begin{pmatrix} 2-i \\ 5 \end{pmatrix}\right\}$

5. $-3, -3$; $E_{-3} =$ gen $\left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right\}$; la mult. geom. es 1

7. $0, 1, 3$; $E_0 =$ gen $\left\{\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}\right\}$; $E_1 =$ gen $\left\{\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}\right\}$; $E_3 =$ gen $\left\{\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}\right\}$

9. $1, 1, 10$; $E_1 =$ gen $\left\{\begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -2 \end{pmatrix}\right\}$; $E_{10} =$ gen $\left\{\begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}\right\}$; la mult. geom. de 1 es 2

11. $1, 1, 1$; $E_1 =$ gen $\left\{\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}\right\}$; la mult. geom. es 1 (la mult. alg. es 3)

13. $-1, i, -i$; $E_{-1} =$ gen $\left\{\begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}\right\}$; $E_i =$ gen $\left\{\begin{pmatrix} 1+i \\ 1 \\ 1 \end{pmatrix}\right\}$; $E_{-i} =$ gen $\left\{\begin{pmatrix} 1-i \\ 1 \\ 1 \end{pmatrix}\right\}$

15. $1, 2, 2$; $E_1 =$ gen $\left\{\begin{pmatrix} 4 \\ 1 \\ -3 \end{pmatrix}\right\}$; $E_2 =$ gen $\left\{\begin{pmatrix} 3 \\ 1 \\ -2 \end{pmatrix}\right\}$; la mult. geom. de 2 es 1

17. a, a, a, a; $E_a = \mathbb{R}^4$; la mult. geom. de a = la mult. alg. de a = 4.

19. a, a, a, a;

$$E_a = \text{gen}\left\{\begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\\1\end{pmatrix}\right\};$$

la mult. alg. de a = 4; la mult. geom. de a = 2.

21. Los valores característicos son $a \pm ib$. Entonces

$$[A-(a+ib)I]\begin{pmatrix}1\\i\end{pmatrix} = \begin{pmatrix}-ib & b\\ -b & -ib\end{pmatrix}\begin{pmatrix}1\\i\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix}.$$

Análogamente,

$$[A-(a-ib)I]\begin{pmatrix}1\\-i\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix}.$$

23. Sean $\beta_1, \beta_2, \ldots, \beta_m$ los valores característicos de αA. Entonces, para cada i, existe un vector $\mathbf{v}_i \neq \mathbf{0}$ tal que $(\alpha A)\mathbf{v}_i = \beta_i \mathbf{v}_i$. Por consiguiente $(\alpha A - \beta_i I)\mathbf{v} = \mathbf{0}$, o bien $\alpha\left(A - \frac{\beta_i}{\alpha}I\right)\mathbf{v} = \mathbf{0}$, lo cual implica que $\det\left(A - \frac{\beta_i}{\alpha}I\right) = 0$.

En consecuencia $\frac{\beta_i}{\alpha}$ es un valor característico de A y $\frac{\beta_i}{\alpha} = \lambda_j$ para alguna j y $\beta_i = \alpha\lambda_j$. Así que cada valor característico de αA es de la forma $\alpha\lambda_j$. Recíprocamente, si $\mu_i = \alpha\lambda_i$, elíjase un vector $\mathbf{v}_i$ tal que $A\mathbf{v}_i = \lambda_i\mathbf{v}_i$. Entonces $(\alpha A)\mathbf{v}_i = \alpha\lambda_i\mathbf{v}_i = \mu_i\mathbf{v}_i$ de modo que μ_i es un valor característico de αA.

25. $\det(A - \lambda_i I) = 0$ y $\det A^{-1} \neq 0$ porque $(A^{-1})^{-1} = A$ existe. Por lo tanto $0 = \det(A - \lambda_i I)\det A^{-1} = \det[(A-\lambda_i I)A^{-1}] = \det(I - \lambda_i A^{-1})$

$$= \det \frac{1}{\lambda_i}\left|\frac{1}{\lambda_i}I - A^{-1}\right|$$

$$= \frac{1}{\lambda_i^n}\det\left|\frac{1}{\lambda_i}I - A^{-1}\right|.$$ En el último paso aplicamos el hecho de que $\det\alpha A = \alpha^n \det A$ si A es una matriz de $n \times n$ (vea el Problema 2.2.28), y $\lambda_i \neq 0$ por el Teorema 6, partes (i) y (x). Consecuentemente

$$\det\left(A^{-1} - \frac{1}{\lambda_i}I\right) = (-1)^n\det\left(\frac{1}{\lambda_i}I - A^{-1}\right) = 0,$$

y $\frac{1}{\lambda_i}$ es un valor característico de A^{-1}. Recíprocamente, si μ_i es un valor característico de A^{-1}, entonces, puesto que $(A^{-1})^{-1} = I, \frac{1}{\mu_i}$ es un valor característico de A de modo que $\frac{1}{\mu_i} = \lambda_j$ para alguna j, o bien $\mu_i = \frac{1}{\lambda_j}$. Por ello todos los valores característicos de A^{-1} son de la forma $\frac{1}{\lambda_i}$. De manera alternativa, si $A\mathbf{v} = \lambda\mathbf{v}$, entonces $\mathbf{v} = A^{-1}\lambda\mathbf{v}$ de forma que $A^{-1}\mathbf{v} = \frac{1}{\lambda}\mathbf{v}$ y $\frac{1}{\lambda}$ es un valor característico de A^{-1}.

Como $\det(A - \lambda_i I) = 0$, vemos que $\det(A^2 - \lambda_i^2 I) = \det(A - \lambda_i I)(A + \lambda_i I) = \det(A - \lambda_i I)\det(A + \lambda_i I) = 0$. Por consiguiente, si λ_i^2 es un valor característico de A^2. Recíprocamente, si μ_i es un valor característico de A^2, entonces $0 = \det(A^2 - \mu_i I) = \det[(A - \sqrt{\mu_i}I)(A + \sqrt{\mu_i}I)]$ de manera que $\det(A - \sqrt{\mu_i}I)$ o bien $\det(A + \sqrt{\mu_i}I) = 0$. En cualquier caso $\pm\sqrt{\mu_i}$ es un valor característico de A de modo que $\mu_i = (\pm\sqrt{\mu_i})^2 = (\pm\lambda_j)^2 = \lambda_j^2$ para alguna j.

29. Del Problema 28, $a_iA^i\mathbf{v} = a_i\lambda^i\mathbf{v}$ por lo que $p(A)\mathbf{v} =$

$$\sum_{i=1}^{n} (a_iA^i)\mathbf{v} = \sum_{i=0}^{n} (a_iA^i\mathbf{v}) = \sum_{i=0}^{n} a_i\lambda^i\mathbf{v} = \left(\sum a_i\lambda^i\right)\mathbf{v} = p(\lambda)\mathbf{v}.$$

31. Si A es triangular superior, entonces también lo es $A - \lambda I$ de forma que, por el Teorema 2.1.1, $\det(A - \lambda I) = (a_{ii} - \lambda)(a_{22} - \lambda)\cdots(a_{nn} - \lambda) = 0$ cuando $\lambda = a_{ii}$ para alguna $i = 1, 2, \ldots, n$.

33. $A\mathbf{v} = \lambda\mathbf{v}$ en donde $\mathbf{v} \neq 0$. Entonces $\overline{A\mathbf{v}} = \overline{\lambda\mathbf{v}}$, lo que implica que $\bar{A}\bar{\mathbf{v}} = \bar{\lambda}\bar{\mathbf{v}}$. Pero si A es real, entonces $\bar{A} = A$. Por lo tanto $A\bar{\mathbf{v}} = \bar{\lambda}\bar{\mathbf{v}}$ y $\mathbf{v} \neq \mathbf{0}$ de manera que $\bar{\lambda}$ es un valor característico de A con vector característico $\bar{\mathbf{v}}$. Aquí hemos utilizado el hecho fácilmente verificable de que $\overline{A\mathbf{v}} = \bar{A}\bar{\mathbf{v}}$.

Problemas 6.2

1.

n	$p_{j,n}$	$p_{a,n}$	T_n	$p_{j,n}/p_{a,n}$	T_n/T_{n-1}
0	0	12	12	0	—
1	36	7	43	5.14	3.58
2	21	19	40	1.11	0.930
5	104	45	149	2.31	—
10	600	291	891	2.06	—
19	16,090	7737	23827	2.08	—
20	23,170	11140	34310	2.08	1.44

Nótese que los valores característicos son 1.44 y −0.836.

Los vectores característicos correspondientes son $\begin{pmatrix}2.08\\1\end{pmatrix}$ y $\begin{pmatrix}-3.57\\1\end{pmatrix}$.

3.

n	$p_{j,n}$	$p_{a,n}$	T_n	$p_{j,n}/p_{a,n}$	T_n/T_{n-1}
0	0	20	20	0	—
1	80	16	96	5	4.8
2	64	69	133	0.928	1.39
5	1092	498	1590	2.19	—
10	3114	1970	5084	1.58	—
19	3.69×10^7	1.95×10^7	5.64×10^7	1.89	—
20	7.82×10^7	4.14×10^7	11.96×10^7	1.89	2.12

Los valores característicos son 2.12 y −1.32 con los vectores característicos correspondientes $\begin{pmatrix}1.89\\1\end{pmatrix}$ y $\begin{pmatrix}-3.03\\1\end{pmatrix}$.

5. De la Ecuación (9), $p_n \approx a_1\lambda_1^n\mathbf{v}_i$ para n grande. Si $\mathbf{v}_1 = \begin{pmatrix}x\\y\end{pmatrix}$, entonces $\frac{p_{j,n}}{p_{a,n}} \approx \frac{a_1\lambda_1^n x}{a_1\lambda_1^n y} = \frac{x}{y}$; pero $\begin{pmatrix}-\lambda_1 & k\\ \alpha & \beta - \lambda_1\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix}$ de forma que $-\lambda_1 x + ky = 0$ y $\frac{x}{y} = \frac{k}{\lambda_1}$. Por lo tanto $\frac{p_{j,n}}{p_{a,n}} \approx \frac{x}{y} = \frac{k}{\lambda_1}$ para n grande.

Problemas 6.3

1. sí; $C = \begin{pmatrix}1 & 2\\ 1 & -5\end{pmatrix}$, $C^{-1}AC = \begin{pmatrix}-4 & 0\\ 0 & 3\end{pmatrix}$

3. sí; $C = \begin{pmatrix}1 & 1\\ 2-i & 2+i\end{pmatrix}$; $C^{-1}AC = \begin{pmatrix}i & 0\\ 0 & -i\end{pmatrix}$

5. sí; $C = \begin{pmatrix}2 & 2\\ -1+3i & -1-3i\end{pmatrix}$; $C^{-1}AC = \begin{pmatrix}2+3i & 0\\ 0 & 2-3i\end{pmatrix}$

7. sí; $C = \begin{pmatrix}3 & 1 & 1\\ 2 & 3 & 0\\ 1 & 1 & 1\end{pmatrix}$; $C^{-1}AC = \begin{pmatrix}1 & 0 & 0\\ 0 & 2 & 0\\ 0 & 0 & -1\end{pmatrix}$

9. sí; $C = \begin{pmatrix}0 & 0 & 1\\ 1 & 1 & 0\\ 0 & 2 & 0\end{pmatrix}$; $C^{-1}AC = \begin{pmatrix}0 & 0 & 0\\ 0 & 2 & 0\\ 0 & 0 & 3\end{pmatrix}$

11. $C = \begin{pmatrix}1 & 0 & 2\\ 3 & -2 & 1\\ 0 & 1 & 2\end{pmatrix}$; $C^{-1}AC = \begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 2\end{pmatrix}$

13. No, dado que 1 es un valor característico de mult. alg. 3 y mult. geom. 1.

15. sí; $C = \begin{pmatrix}0 & -1 & 1 & 1\\ 0 & 1 & 1 & 1\\ 1 & 0 & 1 & -1\\ 0 & 1 & -1 & 1\end{pmatrix}$; $C^{-1}AC = \begin{pmatrix}2 & 0 & 0 & 0\\ 0 & 2 & 0 & 0\\ 0 & 0 & 4 & 0\\ 0 & 0 & 0 & 6\end{pmatrix}$

17. $B = C^{-1}AC$ de modo que $B^n = (C^{-1}AC)(C^{-1}AC)\cdots(C^{-1}AC) = C^{-1}A(CC^{-1})A(CC^{-1})\cdots \times AC = C^{-1}AIA\cdots IAC = C^{-1}AA\cdots AC = C^{-1}A^nC$.

19. $\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$

21. Si $D = C^{-1}AC$, entonces, como en el Problema 17, $D^n = (C^{-1}AC) \times (C^{-1}AC)\cdots(C^{-1}AC) = C^{-1}A^nC$ de modo que $A^n = CD^nC^{-1}$.

23. Claramente A tiene a c como un valor característico de mult. alg. n. Consecuentemente si A es diagonalizable, debe haber una matriz invertible E tal que $E^{-1}AE = \text{diag}(c, c, \ldots, c) = cI$ de manera que $A = E(cI)E^{-1} = cEIE^{-1} = cI$.

25. Si A y B tienen valores característicos distintos, entonces ambos tienen n vectores característicos linealmente independientes, y se tiene que $D_1 = C_1^{-1}AC_1$ y $D_2 = C_2^{-1}BC_2$.

(i) Si A y B tienen los mismos vectores característicos, entonces $C_1 = C_2 = C$ y $AB = (CD_1C^{-1}) \times (CD_2C^{-1}) = CD_1D_2C^{-1} = CD_2D_1C^{-1} = (CD_2C^{-1}) \times (CD_1C^{-1}) = BA$ (ya que las matrices diagonales del mismo orden siempre conmutan).
(ii) Si $BA = AB$, sea $\mathbf{x}$ un vector característico de B que corresponde a λ. Entonces $BA\mathbf{x} = AB\mathbf{x} = A(\lambda\mathbf{x}) = \lambda A\mathbf{x}$ por lo que $y = A\mathbf{x}$ es un vector característico de B correspondiente a λ. Por ello $A\mathbf{x}$ y $\mathbf{x}$ son linealmente dependientes de modo que existe un escalar μ con $A\mathbf{x} = \mu\mathbf{x}$. Pero esto demuestra que $\mathbf{x}$ es también un vector característico de A. Consecuentemente todo vector característico de B es un vector característico de A. Un razonamiento semejante muestra que todo vector característico de A es un vector característico de B.

Problemas 6.4

1. $Q = \begin{pmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ 1/\sqrt{5} & -2/\sqrt{5} \end{pmatrix}$, $D = \begin{pmatrix} 5 & 0 \\ 0 & -5 \end{pmatrix}$

3. $Q = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix}$, $D = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}$

5. $Q = \begin{pmatrix} 1/\sqrt{2} & \frac{1}{2} & \frac{1}{2} \\ -1/\sqrt{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix}$,

$D = \begin{pmatrix} -3 & 0 & 0 \\ 0 & 1+2\sqrt{2} & 0 \\ 0 & 0 & 1-2\sqrt{2} \end{pmatrix}$

7. $Q = \begin{pmatrix} -\frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \end{pmatrix}$, $D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 6 \end{pmatrix}$

9. Sea $\mathbf{u}$ un vector característico que corresponde a λ con $|\mathbf{u}| = 1$. Entonces $Q\mathbf{u} = \lambda\mathbf{u}$ y $1 = |\mathbf{u}| = |Q^{-1}Q\mathbf{u}| = |\lambda Q^{-1}\mathbf{u}| = |\lambda Q^t\mathbf{u}| = |\lambda Q\mathbf{u}| =$ (puesto que Q es simétrica) $|\lambda^2\mathbf{u}| = \lambda^2|\mathbf{u}| = \lambda^2$. Así que $\lambda^2 = 1$ y $\lambda = \pm 1$.

11. $1 = \det I = \det(Q^{-1}Q) = \det(Q^tQ) = (\det Q^t)(\det Q) = (\det Q)^2$ – pues $\det A^t = \det A$ para cualquier matriz A. Por consiguiente $\det Q = \pm 1$ y $\begin{pmatrix} a & c \\ b & d \end{pmatrix} = Q^t =$

$$Q^{-1} = \begin{pmatrix} \dfrac{d}{\det Q} & -\dfrac{b}{\det Q} \\ -\dfrac{c}{\det Q} & \dfrac{a}{\det Q} \end{pmatrix}.$$

Si $\det Q = 1$, entonces $c = -b$. Si $\det Q = -1$, entonces $c = b$.

13. Si la matriz A de 2×2 tiene vectores característicos ortogonales, entonces A es ortogonalmente diagonalizable, lo que significa que A es simétrica por el Teorema 4.

15. Sea λ un valor característico de A con vector característico $\mathbf{v}$ y suponga que $A^* = A$. Entonces $\lambda(\mathbf{v}, \mathbf{v}) = (\lambda\mathbf{v}, \mathbf{v}) = (A\mathbf{v}, \mathbf{v}) = (\mathbf{v}, A^*\mathbf{v}) = (\mathbf{v}, A\mathbf{v}) = (\mathbf{v}, \lambda\mathbf{v}) = \bar{\lambda}(\mathbf{v}, \mathbf{v})$. Como $\mathbf{v} \neq 0$, esto significa que $\lambda = \bar{\lambda}$ de manera que λ es real.

17. Utilice el Problema 16 después de mostrar que a todo valor característico de mult. alg. k corresponden k vectores característicos ortonormales. Sea Q obtenida exactamente como en la demostración del Teorema 3. Recuérdese que $(\mathbf{u}, \mathbf{v}) = \mathbf{u}_1 \cdot \bar{\mathbf{v}}_1 + \cdots + \mathbf{u}_n \cdot \bar{\mathbf{v}}_n$. $Q^t = \bar{Q}^{-1}$ y A es semejante a $Q^tA\bar{Q}$ o $\bar{Q}^tAQ$. $|Q^tA\bar{Q} - I| = |A - I|$; $\bar{Q}^tAQ = (\bar{Q}^tA)Q$

$$= \begin{pmatrix} \bar{\mathbf{u}}_1^tA \\ \bar{\mathbf{u}}_2^tA \\ \vdots \\ \bar{\mathbf{u}}_n^tA \end{pmatrix}(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n)$$

$$= \begin{pmatrix} \bar{\mathbf{u}}_1^t \\ \bar{\mathbf{u}}_1^t \\ \vdots \\ \bar{\mathbf{u}}_n^t \end{pmatrix}(A^*\mathbf{u}_1, A^*\mathbf{u}_2, \ldots, A^*\mathbf{u}_n).$$

Ahora bien $(\bar{\mathbf{u}}_1^t, A^*\mathbf{u}_1) = (\bar{\mathbf{u}}_1^t, A\mathbf{u}_1) = (\bar{\mathbf{u}}_1^t, \lambda_1\mathbf{u}_1) = \bar{\lambda}_1(\bar{\mathbf{u}}_1^t, \mathbf{u}_1) = \bar{\lambda}_1 = \lambda_1$ (por el Problema 15) y dado que $(\bar{\mathbf{u}}_1^t, \mathbf{u}_1) = \bar{\mathbf{u}}_1^t \cdot \bar{\mathbf{u}}_1 = 1 = \overline{\mathbf{u}_1^t \cdot \mathbf{u}_1}$. Entonces $Q_tAQ =$

$$\begin{pmatrix} \lambda_1 & \bar{\mathbf{u}}_1^tA\mathbf{u}_2 & \cdots & \bar{\mathbf{u}}_1^tA\mathbf{u}_n \\ 0 & \bar{\mathbf{u}}_2^tA\mathbf{u}_2 & \cdots & \bar{\mathbf{u}}_2^tA\mathbf{u}_n \\ \vdots & \vdots & & \vdots \\ 0 & \bar{\mathbf{u}}_n^tA\mathbf{u}_2 & \cdots & \bar{\mathbf{u}}_n^tA\mathbf{u}_n \end{pmatrix}$$

$\bar{u}_1^tAu_j = A\bar{u}_1^t \cdot \bar{u}_j = \overline{Au_1^t \cdot u_j} = 0$ si $j \neq 1$. Ahora $\overline{(\bar{Q}^tAQ)^t} = \overline{Q^tA^t(\bar{Q}^t)^t} = \overline{Q^tA^t\bar{Q}} = \bar{Q}^t\bar{A}^tQ = \bar{Q}^tAQ$, ya que $\bar{A}^t = A^* = A$. En consecuencia $\bar{Q}^tAQ$ es hermitiana, lo que significa que los ceros que están en el primer renglón de $\bar{Q}^tAQ$ deben corresponder a los ceros en la primera columna. El resto de la demostración se sigue, como en la del Teorema 3, con Q^t reemplazada por $\bar{Q}^t$.

19. $U = \frac{1}{\sqrt{3}}\begin{pmatrix} -1+i & 1 \\ 1 & 1+i \end{pmatrix}$; $U^*AU = \begin{pmatrix} -1 & 0 \\ 0 & 8 \end{pmatrix}$

Problemas 6.5

1. $\begin{pmatrix} 3 & -1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = 5$;

$$Q = \begin{pmatrix} \dfrac{2}{\sqrt{26-6\sqrt{13}}} & \dfrac{2}{\sqrt{26+6\sqrt{13}}} \\ \dfrac{3-\sqrt{13}}{\sqrt{26-6\sqrt{13}}} & \dfrac{3+\sqrt{13}}{\sqrt{26+6\sqrt{13}}} \end{pmatrix} = \begin{pmatrix} 0.9571 & 0.2898 \\ -0.2898 & 0.9571 \end{pmatrix}; \quad \frac{x'^2}{\left(\dfrac{10}{\sqrt{13}+3}\right)} - \frac{y'^2}{\left(\dfrac{10}{\sqrt{13}-3}\right)} = 1;$$

hipérbola; $\theta = 5.989 = 343°$

3. $\begin{pmatrix} 4 & 2 \\ 2 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = 9$;

$$Q = \begin{pmatrix} \dfrac{5+\sqrt{41}}{\sqrt{82+10\sqrt{41}}} & \dfrac{5-\sqrt{41}}{\sqrt{82-10\sqrt{41}}} \\ \dfrac{4}{\sqrt{82+10\sqrt{41}}} & \dfrac{4}{\sqrt{82-10\sqrt{41}}} \end{pmatrix} = \begin{pmatrix} 0.9436 & -0.3310 \\ 0.3310 & 0.9436 \end{pmatrix}; \quad \frac{x'^2}{\left(\dfrac{18}{\sqrt{41}+3}\right)} - \frac{y'^2}{\left(\dfrac{18}{\sqrt{41}-3}\right)} = 1;$$

hipérbola; $\theta \approx 0.3374 \approx 19.33°$

5. $\begin{pmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = a > 0$;

$Q = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}$; $\dfrac{x'^2}{2a} - \dfrac{y'^2}{2a} = 1$;

hipérbola; $\theta = 7\pi/4 = 315°$.

7. Igual que en el Problema 5 excepto que ahora tenemos una hipérbola con los papeles de x' y y' invertidos; como $a < 0$, se tiene que

$$\frac{y'^2}{(-2a)} - \frac{x'^2}{(-2a)} = 1.$$

9. $\begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = 0$;

$Q = \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}$;

$y'^2 = 0$, lo cual es la ecuación de una recta que pasa por el origen; $\theta = \pi/4 = 45°$.

11. $\begin{pmatrix} 3 & -3 \\ -3 & 5 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = 36$;

$$Q = \begin{pmatrix} \dfrac{1+\sqrt{10}}{\sqrt{20+2\sqrt{10}}} & \dfrac{1-\sqrt{10}}{\sqrt{20-2\sqrt{10}}} \\ \dfrac{3}{\sqrt{20+2\sqrt{10}}} & \dfrac{3}{\sqrt{20-2\sqrt{10}}} \end{pmatrix} = \begin{pmatrix} 0.8112 & -0.5847 \\ 0.5847 & 0.8112 \end{pmatrix}$$

$$\frac{x'^2}{\left(\dfrac{36}{4-\sqrt{10}}\right)} + \frac{y'^2}{\left(\dfrac{36}{4+\sqrt{10}}\right)} = 1;$$

elipse; $\theta \approx 0.6245 \approx 35.78°$

13. $\begin{pmatrix} 6 & \frac{5}{2} \\ \frac{5}{2} & -6 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = -7$;

$Q = \begin{pmatrix} 5/\sqrt{26} & -1/\sqrt{26} \\ 1/\sqrt{26} & 5/\sqrt{26} \end{pmatrix}$;

$\dfrac{y'^2}{(14/13)} - \dfrac{x'^2}{(14/13)} = 1$;

hipérbola; $\theta \approx 1.377 \approx 78.91°$

15. $\begin{pmatrix} 1 & -1 & -1 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix}$;

$Q = \begin{pmatrix} 1/\sqrt{3} & 1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & -1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & 0 & -2/\sqrt{6} \end{pmatrix}$;

$-x'^2 + 2y'^2 + 2z'^2$

17. $\begin{pmatrix} 3 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 4 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix}$;

$Q = \begin{pmatrix} -\frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \end{pmatrix}$;

$3y'^2 + 6z'^2$

19. $\begin{pmatrix} 1 & 1 & 2 & \frac{7}{2} \\ 1 & 1 & 3 & -1 \\ 2 & 3 & 3 & 0 \\ \frac{7}{2} & -1 & 0 & 1 \end{pmatrix}$

21. $\begin{pmatrix} 3 & -\frac{7}{2} & \frac{1}{2} & -1 & \frac{3}{2} \\ -\frac{7}{2} & -2 & -\frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & -\frac{1}{2} & 3 & -2 & -\frac{5}{2} \\ -1 & \frac{1}{2} & -2 & 6 & \frac{1}{2} \\ \frac{3}{2} & 0 & -\frac{5}{2} & \frac{1}{2} & -1 \end{pmatrix}$

23. $\begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} =$

$$\begin{pmatrix} x\cos\theta & -y\,\text{sen}\,\theta \\ x\,\text{sen}\,\theta & +y\cos\theta \end{pmatrix} = \begin{pmatrix} x' \\ y' \end{pmatrix}.$$

Entonces la ecuación cuadrática $ax'^2 + bx'y' + cy'^2$ se convierte en $a(x\cos\theta - y\,\text{sen}\,\theta)^2 + b(x\cos\theta - y\,\text{sen}\,\theta) \times (x\,\text{sen}\,\theta + y\cos\theta) + c(x\,\text{sen}\,\theta + y\cos\theta)^2$; el término producto de cruz es $-2axy(\text{sen}\,\theta\cos\theta + bxy \times [\cos^2\theta - \text{sen}^2\theta] + 2cxy \times \text{sen}\,\theta\cos\theta = xy[-a\,\text{sen}\,2\theta + b\cos 2\theta + c\,\text{sen}\,2\theta] = 0$ de manera que $(c-a)\text{sen}\,2\theta + b\cos 2\theta = 0$ y $\dfrac{a-c}{b} = \dfrac{\cos 2\theta}{\text{sen}\,2\theta} = \cot 2\theta.$

25. Supóngase que $ax^2 + bxy + cy^2$ se convierte en $a'x'^2 + b'x'y' + c'y'^2$ por rotación. Sean

$$A = \begin{pmatrix} a & \frac{a}{2} \\ \frac{b}{2} & c \end{pmatrix} \quad \text{y} \quad A' = \begin{pmatrix} a' & \frac{b'}{2} \\ \frac{b'}{2} & c' \end{pmatrix}.$$

Existe una matriz ortogonal Q tal que $A = QA'Q^t$ y Q es también una matriz de rotación. Por lo tanto $\det A = \det Q \det A' \det Q^t = \det QQ^t \det A' = \det A'$ ya que $QQ^t = I$. Pero $\det A = ac - \dfrac{b^2}{4}$ y $\det A' = a'c' - \dfrac{b'^2}{4}$.

Finalmente, dado que A y A' son semejantes, tienen los mismos valores característicos. Pero la suma de los valores característicos de A es $a + c$, en tanto que la suma de los de A' es $a' + c'$. Así que $a + c = a' + c'$.

27. Sean $\lambda_1, \lambda_2, \ldots, \lambda_n$ los valores característicos de A. Entonces, quitando los términos de producto de cruz, tenemos $F(\mathbf{x}) = F'(\mathbf{x}') = \lambda_1 x_1'^2 + \lambda_2 x_2'^2 + \cdots + \lambda_n x_2'^2$ en donde $\mathbf{x}' = Q^t\mathbf{x}$. Si $\lambda_i \ge 0$ para $i = 1, 2, \ldots, n$, entonces $F'(\mathbf{x}') \ge 0$. Si $F'(\mathbf{x}') \ge 0$ entonces $\lambda_i \ge 0$ dado que, si no, hay un λ_j con $\lambda_j < 0$. Sea $\mathbf{x}^*$ el vector con ceros en todas las posiciones excepto en la j-ésima posición, con un 1 en ésta. Entonces $H(\mathbf{x}^*) = \lambda_j < 0$, lo cual es una contradicción.

29. negativa definida

31. positiva definida

33. indefinida

35. negativa definida

37. (i) Si $\det A \neq 0$, entonces ni λ_1 ni λ_2 son cero. Consecuentemente, con $d = 0$, la Ecuación (22) se transforma en $\lambda_1 x'^2 + \lambda_2 y'^2 = 0$. Si ahora λ_1 y λ_2 son positivos o bien negativos, entonces la ecuación se satisface sólo cuando $x' = 0$ y $y' = 0$. Éstas son las ecuaciones de dos rectas. Si λ_1 y λ_2 tienen signos opuestos, entonces las ecuaciones se convierten en $x' = \pm\sqrt{\dfrac{\lambda_2}{-\lambda_1}}\,y'$, las cuales son de nuevo las ecuaciones de dos rectas. Si $\det A = 0$, entonces una de λ_1 o de λ_2 es cero, y la ecuación se transforma en $x' = 0$ o bien $y' = 0$, cada una de las cuales es la ecuación de una sola recta.

Problemas 6.6

1. no **3.** no **5.** sí **7.** no
9. sí **11.** no **13.** sí
15. I

17. $\begin{pmatrix} 1 & 0 \\ -1 & -\frac{1}{7} \end{pmatrix}$; $J = \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}$

19. (a) Sea $\mathbf{x} \in \mathbb{C}^3$ un vector fijo lo que no es un vector característico. Dado que la mult. geom. del valor característico λ es 1, $\mathbf{x}$ no es un múltiplo de $\mathbf{v}_1$, en donde $\mathbf{v}_1$ es un vector característico. Entonces $\mathbf{x}$ y $\mathbf{y}_1$ son linealmente independientes. Sea $\mathbf{w} = c_1\mathbf{v}_1 + c_2\mathbf{x}$. Supóngase que $\mathbf{w} = (A - \lambda I)\mathbf{x}$. Entonces $A\mathbf{x} - \lambda\mathbf{x} = c_1\mathbf{v}_1 + c_2\mathbf{x}$. Sea $B = A - (c_2 + \lambda)I$ de modo que $B\mathbf{x} = c_1\mathbf{v}_1$. Considere que $c_2 \neq 0$. Entonces $\lambda + c_2$ no es un valor característico pues λ es el único valor característico. Tenemos $\det B = \det[A - (\lambda + c_2)I] \neq 0$. Entonces B^{-1} existe, $\mathbf{x} = c_1 B^{-1}\mathbf{v}_1$, y $\lambda\mathbf{x} = c_1 B^{-1}\lambda\mathbf{v}_1 = c_1 B^{-1}A\mathbf{v}_1$. Puesto que $A = B + (c_2 + \lambda)I$, $\lambda\mathbf{x} = c_1[B^{-1}B + (c_2 + \lambda)B^{-1}]\mathbf{v}_1 = c_1\mathbf{v}_1 + c_1(c_2 + \lambda)B^{-1}\mathbf{v}_1 = c_1\mathbf{v}_1 + (c_2 + \lambda)B^{-1}B\mathbf{x} = c_1\mathbf{v}_1 + c_2\mathbf{x} + \lambda\mathbf{x}$.

Por consiguiente $c_1\mathbf{v}_1 + c_2\mathbf{x} = \mathbf{0}$ y $c_1 = c_2 = 0$ porque $\mathbf{v}_1$ y $\mathbf{x}$ son linealmente independientes. Esto contradice la hipótesis previa de que $c_2 \neq 0$ y $\mathbf{w} = (A - \lambda I)\mathbf{x} = c_1\mathbf{v}_1$. Sea $c_1\mathbf{x} = \mathbf{v}_2$; entonces $(A - \lambda I)\mathbf{v}_2 = \mathbf{v}_1$.

(b) Sea $\mathbf{y} \in \mathbb{C}^3$ con $\mathbf{y}$ un vector característico no perteneciente a A. $\mathbf{y}$ se puede elegir linealmente independiente de $\mathbf{v}_2$ (ya es independiente de $\mathbf{v}_1$) de forma que $\mathbf{z} = d_1\mathbf{v}_2 + d_2\mathbf{y}$ no es un vector característico. Escríbase $\mathbf{z}$ como $\mathbf{z} = (A - \lambda I)\mathbf{y}$. Entonces $A\mathbf{y} - \lambda\mathbf{y} = d_1\mathbf{v}_2 + d_2\mathbf{y}$. Sea $D = A - (d_2 + \lambda)I$ de manera que $D\mathbf{y} = d_1\mathbf{v}_2$ puesto que $A\mathbf{y} - (\lambda I)\mathbf{y} - d_2\mathbf{y} = d_1\mathbf{v}_2$. Supóngase que $d_2 \neq 0$; entonces $d_2 + \lambda$ no es un valor característico. Claramente $\det D \neq 0$, D^{-1} existe, $\mathbf{y} = d_1 D^{-1}\mathbf{v}_2$ y $\lambda\mathbf{y} = d_1 D^{-1}\lambda\mathbf{v}_2$. Entonces $\lambda\mathbf{v}_2 = \mathbf{v}_1 - A\mathbf{v}_2$; $\lambda\mathbf{y} = d_1 D^{-1} \times (\mathbf{v}_1 - A\mathbf{v}_2) = d_1 D^{-1}\mathbf{v}_1 - d_1 D^{-1}A\mathbf{v}_2$. $A = D + (d_2 + \lambda)I$. $\lambda\mathbf{y} = d_1 D^{-1}\mathbf{v}_1 - d_1\mathbf{v}_2 - d_1 d_2 D^{-1}\mathbf{v}_2 - d_1 D^{-1}\lambda\mathbf{v}_2 = d_1 D^{-1}\mathbf{v}_1 - d_1 D^{-1}A\mathbf{v}_2 - d_1\mathbf{v}_2 - d_1 d_2 D^{-1}\mathbf{v}_2 = d_1 D^{-1}\lambda\mathbf{v}_2 - d_1(\mathbf{v}_2 + d_2 D^{-1}\mathbf{v}_2) = \lambda\mathbf{y} - d_1(I + d_2 D^{-1})\mathbf{v}_2$. Así que $0 = d_1 (I + d_2 D^{-1})\mathbf{v}_2$. $d_1 \neq 0$, de otro modo $\lambda + d_2$ sería un valor característico y $\mathbf{y}$ un vector característico. Entonces $(d_2 DD^{-1} + D)\mathbf{v}_2 = D\mathbf{0} = \mathbf{0}$. $d_2\mathbf{v}_2 + [A - (d_2 + \lambda)I]\mathbf{v}_2 = \mathbf{0}$. $d_2\mathbf{v}_2 + (A - \lambda I)\mathbf{v}_2 - d_2\mathbf{v}_2 = \mathbf{0}$, o bien $(A - \lambda I)\mathbf{v}_2 = \mathbf{0}$, contrario al resultado de la parte (a). En consecuencia $d_2 = 0$ y $(A - I\lambda)\mathbf{y} = d_1\mathbf{v}_2$. Sea $\mathbf{y} = d_1\mathbf{v}_3$; entonces $(A - I\lambda)\mathbf{v}_3 = \mathbf{v}_2$.

(c) Sea $C = (\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ en donde $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ son como antes y linealmente independientes; entonces C^{-1} existe. $AC = A(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) = (A\mathbf{v}_1, A\mathbf{v}_2, A\mathbf{v}_3) = (\lambda\mathbf{v}_1, \mathbf{v}_1 + \lambda\mathbf{v}_2, \mathbf{v}_2 + \lambda\mathbf{v}_3)$;

$$CJ = (\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)\begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix} = (\lambda\mathbf{v}_1, \mathbf{v}_1 + \lambda\mathbf{v}_2, \mathbf{v}_2 + \lambda\mathbf{v}_3) = AC;$$

así que $J = C^{-1}AC$.

21. $C = \begin{pmatrix} 1 & 1 & 0 \\ 0 & -1 & -2 \\ -1 & 0 & -3 \end{pmatrix}$; $J = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$

Si $m = n$, entonces $A^m = 0$ por la definición de índice de nulipotencia. Si $m > n$, entonces $A^m = A^{m-n}A^n = A^{m-n}0 = 0$.

25. $\begin{pmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \lambda_3 & 0 \\ 0 & 0 & 0 & \lambda_4 \end{pmatrix}$, $\begin{pmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & \lambda_2 & 0 \\ 0 & 0 & 0 & \lambda_3 \end{pmatrix}$, $\begin{pmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 1 & 0 \\ 0 & 0 & \lambda_1 & 0 \\ 0 & 0 & 0 & \lambda_2 \end{pmatrix}$, $\begin{pmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & \lambda_2 & 1 \\ 0 & 0 & 0 & \lambda_2 \end{pmatrix}$, $\begin{pmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 1 & 0 \\ 0 & 0 & \lambda_1 & 1 \\ 0 & 0 & 0 & \lambda_1 \end{pmatrix}$.

Aquí los elementos λ_i no son necesariamente distintos. Asimismo, los grupos pueden permutarse sobre la diagonal.

27. $\begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & -4 \end{pmatrix}$ $\begin{pmatrix} 3 & 1 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & -4 \end{pmatrix}$ $\begin{pmatrix} 3 & 1 & 0 & 0 \\ 0 & 3 & 1 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & -4 \end{pmatrix}$

Los grupos de Jordan pueden permutarse a lo largo de la diagonal.

29. $\begin{pmatrix} 4 & 0 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & -3 & 0 \\ 0 & 0 & 0 & 0 & -3 \end{pmatrix}$ $\begin{pmatrix} 4 & 1 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & -3 & 0 \\ 0 & 0 & 0 & 0 & -3 \end{pmatrix}$ $\begin{pmatrix} 4 & 1 & 0 & 0 & 0 \\ 0 & 4 & 1 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & -3 & 0 \\ 0 & 0 & 0 & 0 & -3 \end{pmatrix}$ $\begin{pmatrix} 4 & 1 & 0 & 0 & 0 \\ 0 & 4 & 1 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & -3 & 1 \\ 0 & 0 & 0 & 0 & -3 \end{pmatrix}$ $\begin{pmatrix} 4 & 1 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & -3 & 1 \\ 0 & 0 & 0 & 0 & -3 \end{pmatrix}$ $\begin{pmatrix} 4 & 0 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & -3 & 1 \\ 0 & 0 & 0 & 0 & -3 \end{pmatrix}$

Los grupos de Jordan pueden permutarse a lo largo de la diagonal.

31. $\begin{pmatrix} -7 & 0 & 0 & 0 & 0 \\ 0 & -7 & 0 & 0 & 0 \\ 0 & 0 & -7 & 0 & 0 \\ 0 & 0 & 0 & -7 & 0 \\ 0 & 0 & 0 & 0 & -7 \end{pmatrix}$ $\begin{pmatrix} -7 & 1 & 0 & 0 & 0 \\ 0 & -7 & 0 & 0 & 0 \\ 0 & 0 & -7 & 0 & 0 \\ 0 & 0 & 0 & -7 & 0 \\ 0 & 0 & 0 & 0 & -7 \end{pmatrix}$ $\begin{pmatrix} -7 & 1 & 0 & 0 & 0 \\ 0 & -7 & 1 & 0 & 0 \\ 0 & 0 & -7 & 0 & 0 \\ 0 & 0 & 0 & -7 & 0 \\ 0 & 0 & 0 & 0 & -7 \end{pmatrix}$

$$\begin{pmatrix}-7&1&0&0&0\\0&-7&1&0&0\\0&0&-7&1&0\\0&0&0&-7&0\\0&0&0&0&-7\end{pmatrix}$$

$$\begin{pmatrix}-7&1&0&0&0\\0&-7&1&0&0\\0&0&-7&1&0\\0&0&0&-7&1\\0&0&0&0&-7\end{pmatrix}$$

Los grupos de Jordan pueden permutarse a lo largo de la diagonal.

Problemas 6.7

1. $\frac{1}{7}\begin{pmatrix}5e^{-4t}+2e^{3t} & 2e^{-4t}-2e^{3t}\\ 5e^{-4t}-5e^{3t} & 2e^{-4t}+5e^{3t}\end{pmatrix}$

3. $\begin{pmatrix}2\,\text{sen}\,t+\cos t & -\text{sen}\,t\\ 5\,\text{sen}\,t & -2\,\text{sen}\,t+\cos t\end{pmatrix}$

5. $e^{-3t}\begin{pmatrix}1-7t & -7t\\ 7t & 1+7t\end{pmatrix}$

7. $e^{-5t}\begin{pmatrix}1-7t & 7t\\ -7t & 1+7t\end{pmatrix}$

9. $\begin{pmatrix}4e^t-3e^{2t}+6te^{2t} & -12e^t+12e^{2t}-6te^{2t} & 6te^{2t}\\ e^t-e^{2t}+2te^{2t} & -3e^t+4e^{2t}-2te^{2t} & 2te^{2t}\\ -3e^t+3e^{2t}-4te^{2t} & 9e^t-9e^{2t}+4te^{2t} & -4te^{2t}+e^{2t}\end{pmatrix}$

11. $\mathbf{x}(t) = -\frac{1}{3}\begin{pmatrix}-e^t-2e^{4t} & -e^t+e^{4t}\\ -2e^t+2e^{4t} & -2e^t-e^{4t}\end{pmatrix}$

$\times\begin{pmatrix}x_1(0)\\x_2(0)\end{pmatrix}$, lo cual conduce a

$x_1(t) = \frac{1}{3}[(x_1(0)+x_2(0))e^t + (2x_1(0)-x_2(0))e^{4t}]$

$= \frac{1}{3}[(x_1(0)+x_2(0)) + (2x_1(0)-x_2(0))e^{3t}]e^t$

Si $2x_1(0) < x_2(0)$, entonces la primera población se extinguirá cuando $x_1(0)+x_2(0) = [x_2(0) - 2x_1(0)]e^{3t}$, o bien

$t = \frac{1}{3}\ln\left(\frac{x_1(0)+x_2(0)}{x_2(0)-2x_1(0)}\right)$.

13. $\begin{pmatrix}x_1\\x_2\end{pmatrix}' = \frac{1}{1000}\begin{pmatrix}-30 & 10\\ 30 & -30\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}$;

$\begin{pmatrix}x_1\\x_2\end{pmatrix} = \begin{pmatrix}500\,(e^{\alpha t}+e^{\beta t})\\ 50\sqrt{0.003}\,(e^{\alpha t}-e^{\beta t})\end{pmatrix}$

en donde $\alpha = -0.03+\sqrt{0.0003}\approx-0.0127$ y $\beta=-0.03-\sqrt{0.0003}\approx -0.0473$

15. (a) $\begin{pmatrix}x_1'\\x_2'\end{pmatrix} = \begin{pmatrix}0 & 1\\ -b & -a\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}$

(b) $\det\begin{pmatrix}-\lambda & 1\\ -b & -a-\lambda\end{pmatrix} = \lambda^2 + a\lambda + b\lambda$

de forma que

$p(\lambda) = \lambda^2 + a\lambda + b = 0.$

17. $(1+5t)e^{-3t}$ 19. $\frac{8}{7}e^{5t}+\frac{13}{7}e^{-2t}$

21. Por el Problema 20, $N_3^k = 0$ para $k \geq 3$. Por tanto $e^{N_3 t} = I + N_3 t +$

$N_3^2\frac{t^2}{2} = \begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}$

$+\begin{pmatrix}0&t&0\\0&0&t\\0&0&0\end{pmatrix}+\begin{pmatrix}0&0&t_2^2\\0&0&0\\0&0&0\end{pmatrix}$

$=\begin{pmatrix}1&t&t_2^2\\0&1&0\\0&0&1\end{pmatrix}$.

23. De 6.6.20, $C = \begin{pmatrix}1&1&0\\1&2&1\\0&1&0\end{pmatrix}$ y $J = \begin{pmatrix}-1&1&0\\0&-1&1\\0&0&-1\end{pmatrix}$ de manera que $e^{At} = C^{-1}e^{Jt}C$

$= \begin{pmatrix}1&0&-1\\0&0&1\\-1&1&-1\end{pmatrix}e^{-t}$

$\begin{pmatrix}1&t&t^2/2\\0&1&t\\0&0&1\end{pmatrix}$

$\begin{pmatrix}1&1&0\\1&2&1\\0&1&0\end{pmatrix}$

$= e^{-t}\begin{pmatrix}1+t & 2t+t^2/2 & t\\ 0 & 0 & 0\\ -t & -t-t^2/2 & -t+1\end{pmatrix}$

25. $e^{\lambda t}\begin{pmatrix}0&t&t^2/2&t^3/2\\0&1&t&t^2/2\\0&0&1&t\\0&0&0&1\end{pmatrix}$

27. $\begin{pmatrix}e^{-4t} & te^{-4t} & (t^2/2)e^{-4t} & 0\\ 0 & e^{-4t} & te^{-4t} & 0\\ 0&0&e^{-4t}&0\\ 0&0&0&e^{3t}\end{pmatrix}$

Problemas 6.8

1. (a) $p(\lambda)=\lambda^2+\lambda-12=0$;
(b) $p(A)=A^2+A-12I$

$=\begin{pmatrix}14&2\\5&11\end{pmatrix}$

$+\begin{pmatrix}-2&-2\\-5&1\end{pmatrix}+\begin{pmatrix}-12&0\\0&-12\end{pmatrix}$

$=\begin{pmatrix}0&0\\0&0\end{pmatrix}$

(c) $A^{-1}=\frac{1}{12}\begin{pmatrix}-1&-2\\-5&2\end{pmatrix}$

3. (a) $p(\lambda)=-\lambda^3+4\lambda^2-3\lambda$;
(b) $p(A)=-A^3+4A^2-3A$

$=-\begin{pmatrix}5&-9&4\\-9&18&-9\\4&-9&5\end{pmatrix}$

$+\begin{pmatrix}8&-12&4\\-12&24&-12\\4&-12&8\end{pmatrix}$

$-\begin{pmatrix}3&-3&0\\-3&6&-3\\0&-3&3\end{pmatrix}$

$=\begin{pmatrix}0&0&0\\0&0&0\\0&0&0\end{pmatrix}$

(c) A^{-1} no existe

5. (a) $p(\lambda)=-\lambda^3+3\lambda^2-3\lambda+1=0$

(b) $p(A)=-A^3+3A^2-3A+I$

$=-\begin{pmatrix}1&-3&3\\3&-8&6\\6&-15&10\end{pmatrix}$

$+\begin{pmatrix}0&0&3\\3&-9&9\\9&-24&18\end{pmatrix}$

$-\begin{pmatrix}0&3&0\\0&0&3\\3&-9&9\end{pmatrix}$

$+\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}$

$=\begin{pmatrix}0&0&0\\0&0&0\\0&0&0\end{pmatrix}$

(c) $A^{-1}=\begin{pmatrix}3&-3&1\\1&0&0\\0&1&0\end{pmatrix}$

7. (a) $p(\lambda)=-\lambda^3+6\lambda^2+18\lambda+9=0$
(b) $p(A)=-A^3+6A^2-18A+9I$

$=-\begin{pmatrix}63&54&108\\180&189&324\\168&204&315\end{pmatrix}$

$+\begin{pmatrix}18&72&54\\108&162&216\\150&114&252\end{pmatrix}$

$-\begin{pmatrix}36&18&54\\72&18&108\\18&90&54\end{pmatrix}$

$+\begin{pmatrix}9&0&0\\0&9&0\\0&0&9\end{pmatrix}$

$=\begin{pmatrix}0&0&0\\0&0&0\\0&0&0\end{pmatrix}$

(c) $A^{-1}=\frac{1}{9}\begin{pmatrix}-27&18&-9\\-6&3&0\\19&-11&6\end{pmatrix}$

9. (a) $p(\lambda)=(a-\lambda)^4$
(b) $p(A)=(aI-A)^4$

$=\begin{pmatrix}0&-b&0&0\\0&0&-c&0\\0&0&0&-d\\0&0&0&0\end{pmatrix}^4$

$=\begin{pmatrix}0&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{pmatrix}$

(c) A^{-1}

$=\begin{pmatrix}1/a & -b/a^2 & cb/a^3 & -bcd/a^4\\ 0 & 1/a & -c/a^2 & cd/a^3\\ 0&0&1/a&-d/a^2\\ 0&0&0&1/a\end{pmatrix}$

11. $|\lambda|\leq 7$ y $\text{Re}\,\lambda \geq \frac{13}{6}$

13. $|\lambda|\leq 10.5$ y $-10.5\leq \text{Re}\,\lambda \leq 5.75$

15. Dado que A es simétrica los valores característicos de A son reales. Entonces, por el Teorema de Gershgorin $\lambda = \text{Re}\,\lambda \geq 4-(2+1+\frac{1}{4})=\frac{3}{4}$.

17. (a) $F(\lambda)=B_0C_0+B_0C_1\lambda + B_1C_0\lambda + B_1C_1\lambda^2$

(b) $P(A)Q(A) = (B_0 + B_1A) \times (C_0 + C_1A) = B_0C_0 + B_0C_1A + B_1AC_0 + B_1AC_1A$

$F(A) = B_0C_0 + B_0C_1A + B_1C_0A + B_1C_1A^2$

$F(A) = P(A)Q(A)$ si y sólo si $C_0A = AC_0$ (en el tercer término) y $AC_1A = C_1A^2$ en el cuarto término.

19. $\det A = \lambda_1\lambda_2\cdots\lambda_n$. Si $\det A = 0$, entonces $\lambda_i = 0$ para alguna i. Pero $|\lambda_i - a_{ii}| \leq r_i$ por lo que $|0-a_{ii}| = |a_{ii}| \leq r_i$, lo cual es imposible pues A es en rigor diagonalmente dominante. Por consiguiente $\lambda_i \neq 0$ para $i = 1, 2, \ldots, n$ y $\det A \neq 0$.

Ejercicios de repaso • Capítulo 6

1. $4, -2$;
$E_4 = \text{gen}\left\{\begin{pmatrix}1\\1\end{pmatrix}\right\}$;
$E_{-2} = \text{gen}\left\{\begin{pmatrix}2\\1\end{pmatrix}\right\}$

3. $1, 7, -5$;
$E_1 = \text{gen}\left\{\begin{pmatrix}-6\\3\\4\end{pmatrix}\right\}$;
$E_7 = \text{gen}\left\{\begin{pmatrix}0\\3\\1\end{pmatrix}\right\}$;
$E_{-5} = \text{gen}\left\{\begin{pmatrix}0\\0\\1\end{pmatrix}\right\}$

5. $1, 3, 3+\sqrt{2}i, 3-\sqrt{2}i$;
$E_1 = \text{gen}\left\{\begin{pmatrix}1\\2\\0\\0\end{pmatrix}\right\}$;
$E_3 = \text{gen}\left\{\begin{pmatrix}1\\1\\0\\0\end{pmatrix}\right\}$;
$E_{3+\sqrt{2}i} = \text{gen}\left\{\begin{pmatrix}0\\0\\-1\\\sqrt{2}i\end{pmatrix}\right\}$;
$E_{3-\sqrt{2}i} = \text{gen}\left\{\begin{pmatrix}0\\0\\1\\\sqrt{2}i\end{pmatrix}\right\}$

7. $C=\begin{pmatrix}-3&1\\4&-1\end{pmatrix}$;
$C^{-1}AC=\begin{pmatrix}2&0\\0&-3\end{pmatrix}$

9. $C=\begin{pmatrix}0&-1-i&-1+i\\1&1&1\\-1&1&1\end{pmatrix}$;

$$C^{-1}AC=\begin{pmatrix}-1&0&0\\0&i&0\\0&0&-i\end{pmatrix}$$

11. es no diagonalizable

13. $Q=\begin{pmatrix}1/\sqrt{2}&0&1/\sqrt{2}\\1/\sqrt{2}&0&-1/\sqrt{2}\\0&1&0\end{pmatrix}$;

$$Q^tAQ=\begin{pmatrix}4&0&0\\0&-3&0\\0&0&0\end{pmatrix}$$

15. $C=\begin{pmatrix}1&1&1&-1\\-1&0&0&0\\0&1&0&0\\-1&-1&-1&2\end{pmatrix}$;

$$C^{-1}AC=\begin{pmatrix}-1&0&0&0\\0&-1&0&0\\0&0&3&0\\0&0&0&3\end{pmatrix}$$

17. $\dfrac{x'^2}{8/(3+\sqrt{2})}+\dfrac{y'^2}{8/(3-\sqrt{2})}=1$: elipse

19. $\dfrac{x'^2}{10/(\sqrt{13}+3)}-\dfrac{y'^2}{10/(\sqrt{13}-3)}=1$: hipérbola

21. $4x'^2-3y'^2$

23. $C=\begin{pmatrix}2&-1\\1&0\end{pmatrix}$;

$$C^{-1}AC=\begin{pmatrix}-2&1\\0&-2\end{pmatrix}$$

25. $\begin{pmatrix}-e^t+2e^{-t}&2e^t-2e^{-t}\\-e^t+e^{-t}&2e^t-e^{-t}\end{pmatrix}$

27. $e^{-t}\begin{pmatrix}\cos 2t-\operatorname{sen}2t&-2\operatorname{sen}2t\\\operatorname{sen}2t&\cos 2t+\operatorname{sen}2t\end{pmatrix}$

29. $|\lambda|\leq 5$ y $-5\leq \operatorname{Re}\lambda\leq \frac{14}{3}$

Capítulo 7

Problemas 7.1

1. 0.33333333×10^0

3. -0.35×10^{-4}

5. 0.77777777×10^0

7. 0.77272727×10^1

9. -0.18833333×10^2

11. 0.23705963×10^9

13. 0.83742×10^{-20}

15. $\varepsilon_a=0.1$, $\varepsilon_r=0.0002$

17. $\varepsilon_a=0.005$, $\varepsilon_r=0.04$

19. $\varepsilon_a=0.00333\ldots$, $\varepsilon_r\approx 0.57143\times 10^{-3}$

21. $\varepsilon_a=1$, $\varepsilon_r\approx 0.1419144\times 10^{-4}$

Problemas 7.2

1. $x_1=1.5$, $x_2=-0.800002$ (el valor real es -0.8), $x_3=-3.7$.

3. $x_1=-0.000001$, $x_2=2.61001$, $x_3=4.3$. La solución exacta es $(0, -2.6.1, 4.3)$.

5. (a) con pivoteo: $x_1=5.99$, $x_2=-2$, $x_3=3.99$
(b) sin pivoteo: $x_1=6$, $x_2=-2$ y $x_3=4$ (Sí, a veces es mejor seguir la ruta más simple. En el Problema 6, con el pivoteo se obtienen respuestas mucho más exactas.) Los errores relativos con pivoteo son $\frac{1}{600}=0.0017$, 0, y $\frac{1}{400}=0.0025$.

7. Una solución con redondeo a 3 cifras significativas es $x_1=1050$ y $x_2=-1000$. La solución exacta es $\mathbf{x}_1=\frac{15650}{13}\approx 1204$ y $\mathbf{x}_2=-\frac{15000}{13}\approx -1154$ Los errores relativos son $0.1465\approx 15\%$ y $0.1333\approx 13\%$.

Problemas 7.3

1. sí **3.** no **5.** sí

7. Jacobi: $x_1=2.9757$, $x_2=-0.9919$ (8 iteraciones) Gauss-Seidel: $x_1=3.0041$, $x_2=-1.0025$, (5 iteraciones). La solución exacta es $(3, -1)$.

9. Jacobi: $x_1=1.9999$, $x_2=-2.988$, $x_3=7.0146$ (7 iteraciones). Gauss-Seidel: $x_1=-1.9975$, $x_2=-2.9993$, $x_3=6.999$ (6 iteraciones). La solución exacta es $(-2, -3, 7)$.

11. Jacobi: $x_1=-8.2863$, $x_2=14.386$, $x_3=-0.10281$ (13 iteraciones) Gauss-Seidel: $x_1=-8.2989$, $x_2=14.399$, $x_3=-0.10025$ (7 iteraciones). La solución exacta es $(-8.3, 14.4, -0.1)$.

13. (a) $|a_{ii}|=1$ y
$$|a_{12}|+|a_{13}|=|a_{21}|+|a_{23}|=|a_{31}|+|a_{32}|=1;$$
por lo tanto $|a_{ii}|=\sum_{\substack{j=1\\j\neq i}}^{3}|a_{ij}|$

de modo que $a_{ii}\ngtr\sum_{\substack{j=1\\j\neq i}}^{3}|a_{ij}|$

(b) *Jacobi*

n	$x_1^{(n)}$	$x_2^{(n)}$	$x_3^{(n)}$
0	0.8	0.8	0.8
1	1.2	1.2	1.2
2	0.8	0.8	0.8
3	1.2	1.2	1.2
4	0.8	0.8	0.8
5	1.2	1.2	1.2

(c) *Gauss–Seidel*

n	$x_1^{(n)}$	$x_2^{(n)}$	$x_3^{(n)}$
0	0.8	0.8	0.8
1	1.2	1	0.9
2	1.05	1.025	0.9625
3	1.0063	1.0156	0.98905
4	0.99768	1.0066	0.99786
5	0.99777	1.0022	1
6	0.99890	1.0006	1.0003
7	0.99955	1.0001	1.0002
8	0.99985	0.99998	1.0001

(d) (i) $D^{-1}(L+U)$
$$=\begin{pmatrix}0&\frac{1}{2}&\frac{1}{2}\\\frac{1}{2}&0&\frac{1}{2}\\\frac{1}{2}&\frac{1}{2}&0\end{pmatrix},$$ la cual tiene el valor característico 1; así por el Teorema 2(i), las iteraciones de Jacobi no convergen.

(ii) $(D+L)^{-1}U$
$$=\begin{pmatrix}-\frac{3}{8}&-\frac{1}{8}&0\\\frac{1}{4}&-\frac{1}{4}&0\\\frac{1}{2}&\frac{1}{2}&0\end{pmatrix},$$
cuyo polinomio característico es $-\lambda(\lambda^2+\frac{5}{8}\lambda+\frac{1}{8})=0$ con raíces 0, $(5\pm\sqrt{7}i)/16$.

$$|(5+\sqrt{7}i)/16|=|(5-\sqrt{7}i)/16|=\sqrt{(\tfrac{5}{16})^2+(\sqrt{7}/16)^2}\approx 0.3536.$$

En consecuencia, $r[(D+L)^{-1}U]\approx 0.35<1$ y, por el Teorema 2(ii), las iteraciones de Gauss-Seidel convergen.

15. Si A es diagonal, entonces $L=U=0$. Entonces $r[D^{-1}(L+U)]=r(0)=0$, y $r[(D+L)^{-1}U]=r(D0)=r(0)=0$.

17. $D=\begin{pmatrix}a&0\\0&d\end{pmatrix}$, $L=\begin{pmatrix}0&b\\0&0\end{pmatrix}$, y

$U=\begin{pmatrix}0&0\\c&0\end{pmatrix}$ de modo que

$$D^{-1}(L+U)=\begin{pmatrix}0&\frac{b}{a}\\\frac{c}{a}&0\end{pmatrix}$$ y

$$(D+L)^{-1}U=\begin{pmatrix}-\frac{bc}{ad}&0\\\frac{c}{d}&0\end{pmatrix}.$$

Los valores característicos de $D^{-1}(L+U)$ son $\pm\sqrt{\dfrac{bc}{ad}}$, y los de $(D+L)^{-1}U$ son 0 y $-\dfrac{bc}{ad}$. Estos son todos menores que uno en valor absoluto si y sólo si $\left|\dfrac{bc}{ad}\right|<1$.

19. $D^{-1}(L+U)$
$$=\begin{pmatrix}\frac{1}{a_{11}}&0&\cdots&0\\0&\frac{1}{a_{22}}&\cdots&0\\\vdots&\vdots&&\vdots\\0&0&\cdots&\frac{1}{a_{nn}}\end{pmatrix}\times\begin{pmatrix}0&a_{12}&a_{13}&\cdots&a_{1n}\\a_{21}&0&a_{23}&\cdots&a_{2n}\\\vdots&\vdots&\vdots&&\vdots\\a_{n1}&a_{n2}&a_{n3}&\cdots&0\end{pmatrix}$$
$$=\begin{pmatrix}0&\frac{a_{12}}{a_{11}}&\frac{a_{13}}{a_{11}}&\cdots&\frac{a_{1n}}{a_{11}}\\\frac{a_{21}}{a_{22}}&0&\frac{a_{23}}{a_{22}}&\cdots&\frac{a_{2n}}{a_{22}}\\\vdots&\vdots&\vdots&&\vdots\\\frac{a_{n1}}{a_{nn}}&\frac{a_{n2}}{a_{nn}}&\frac{a_{n3}}{a_{nn}}&\cdots&0\end{pmatrix}.$$

Como A es en forma estricta diagonalmente dominante, $|a_{11}|>|a_{12}|+|a_{13}|+\cdots+|a_{1n}|$. Así, $\left|\dfrac{a_{12}}{a_{11}}\right|+\left|\dfrac{a_{13}}{a_{11}}\right|+\left|\dfrac{a_{1n}}{a_{11}}\right|<1$.

Un hecho similar se verifica en todo renglón de $D^{-1}(L+U)$. Por tanto, por el Teorema de Gershgorin, todos los valores característicos de $D^{-1}(L+U)$ están situados en círculos centrados en el origen con radios menores que 1. Esto significa que si λ es un valor característico de $D^{-1}(L+U)$, entonces $|\lambda|<1$, lo cual significa que $r[D^{-1}(L+U)]<1$.

Problemas 7.4

1. -4, $\begin{pmatrix}1\\1\end{pmatrix}$ **3.** -6.3, $\begin{pmatrix}1\\-0.5\end{pmatrix}$

5. 8, $\begin{pmatrix}1\\0.5\\1\end{pmatrix}$

7. (a) $\lambda\approx 8.5559$ (sin normalización); $\varepsilon_r\approx 0.00051427$
(b) $\lambda=2+\sqrt{43}=8.5574$. El error relativo real es 0.00017979.

9. Comenzando con $\mathbf{x}_0=\begin{pmatrix}1\\1\end{pmatrix}$ y sin normalización, obtenemos $\begin{pmatrix}1\\1\end{pmatrix}$, $\begin{pmatrix}2\\1\end{pmatrix}$, $\begin{pmatrix}-1\\-1\end{pmatrix}$, $\begin{pmatrix}-2\\-1\end{pmatrix}$, $\begin{pmatrix}1\\1\end{pmatrix}$, $\begin{pmatrix}2\\1\end{pmatrix}$, ... Esto es, el método no converge. Nótese que los valores característicos de A son $\pm i$, ambos con un valor absoluto de 1. Esto es, A no tiene un valor característico dominante.

11. El segundo valor característico es $\lambda_2=3$.

13. Los otros valores característicos son -2 y 1.

Ejercicios de repaso • Capítulo 7

1. $\varepsilon_a=0.02$, $\varepsilon_r=0.002857$

3. $\varepsilon_a=\frac{1}{300}=0.003333\ldots$, $\varepsilon_r=0.25$

5. $\begin{pmatrix}1&-2&3\\0&1&-2\\0&0&1\end{pmatrix}$

7. $x_1=7.11004$, $x_2=-2.39005$, $x_3=-5.92009$ La solución exacta es $(7.11, -2.39, 5.92)$.

9. no

11. Jacobi: $x_1=-0.34008$, $x_2=0.260018$, $x_3=0.509983$ (12 iteraciones) Gauss-Seidel: $x_1=-0.340001$, $x_2=0.26$, $x_3=0.51$ (7 iteraciones). La solución exacta es $(-0.34, 0.26, 0.51)$.

13. $-8, \begin{pmatrix} 1 \\ -\frac{2}{3} \end{pmatrix}$ **15.** $\lambda_2=3$

Apéndice 1

1. Si $n=1$, $1^3=\dfrac{1^2(1+1)^2}{4}=\dfrac{1.4}{4}=1$ de modo que el resultado es válido para $n = 1$. Supóngase que se verifica para $n = k$. Entonces $1^3 + 2^3 + 3^3 + \cdots + k^3 + (k + 1)^3$

$$=\frac{k^2(k+1)^2}{4}+(k+1)^3$$

$$=(k+1)^2\left[\frac{k^2}{4}+(k+1)\right]$$

$$=\frac{(k+1)^2}{4}[k^2+4(k+1)]$$

$$=\frac{(k+1)^2(k^2+4k+4)}{4}$$

$$=\frac{(k+1)^2(k+2)^2}{4},$$

lo que constituye el resultado para $n = k + 1$.

3. Para $n=1$, $\dfrac{1-(1+1)a^1+1a^2}{(1-a)^2}$

$$=\frac{1-2a+a^2}{(1-a)^2}=\frac{(1-a)^2}{(1-a)^2}=1$$

así que el resultado se cumple para $n = 1$. Supóngase que es válido para $n = k$. Entonces

$1+2a+3a^2+\cdots+ka^{k-1}+(k+1)a^k$

$$=\frac{1-(k+1)a^k+ka^{k+1}}{(1-a)^2}+(k+1)a^k$$

$$=\frac{1-(k+1)a^k+ka^{k+1}+(k+1)a^k(1-a)^2}{(1-a)^2}$$

$$=\frac{1-(k+1)a^k+ka^{k+1}+(k+1)a^k(1-2a+a^2)}{(1-a)^2}$$

$$=\frac{1-(k+1)a^k+ka^{k+1}+(k+1)a^k-2(k+1)a^{k+1}+(k+1)a^{k+2}}{(1-a)^2}$$

$$=\frac{1-(k+2)a^{k+1}+(k+1)a^{k+2}}{(1-a)^2},$$

lo que es el resultado para $n = k + 1$.

5. Un polinomio de primer grado tiene la forma $p(x) = ax + b$ con $a \neq 0$ y $x = -b/a$ es la solución única de $p(x) = 0$. Por consiguiente, un polinomio de primer grado tiene exactamente una raíz, y se demuestra el resultado general utilizando inducción matemática sobre el grado del polinomio. El resultado es cierto para $n = 1$, y lo suponemos verdadero para $n = k$. Sea $p(x)$ un polinomio de grado $k + 1$. Por hipótesis, $p(x) = 0$ tiene por lo menos una solución x_0. Entonces $p(x) = (x - x_0)q(x)$ en donde $q(x)$ es un polinomio de grado k. Por la hipótesis de inducción $q(x)$ tiene k raíces $x_1, x_2, \ldots, x_k$. Es decir, $q(x) = (x - x_1)(x - x_2) \cdots (x - x_n)$. Pero entonces, evidentemente, $x_0, x_1, x_2, \ldots, x_k$ son $k + 1$ soluciones de la ecuación $p(x) = 0$. En consecuencia, $p(x)$ tiene $k + 1$ raíces y el resultado se demuestra para $n = k + 1$.

7. Ya que $(A+B)^t=A^t+B^t$, el resultado es cierto para $n = 2$. Suponemos que se cumple para $n = k$. Entonces

$(A_1+A_2+\cdots+A_k+A_{k+1})^t$

$=(A_1+A_2+\cdots+A_k)^t+A_{k+1}^t$ (usando el caso $n=2$)

$=(A_1^t+A_2^t+\cdots+A_k^t)+A_{k+1}^t$ (utilizando el caso $n = k$)

$=A_1^t+A_2^t+\cdots+A_k^t+A_{k+1}^t$

lo que es el resultado para $n = k + 1$.

Apéndice 2

1. $9-7i$ **3.** 2 **5.** $-27+5i$

7. $5\sqrt{2}e^{i(\pi/4)}$

9. $3\sqrt{2}e^{i(7\pi/4)}=3\sqrt{2}e^{-i(\pi/4)}$

11. $6e^{i(\pi/6)}$

13. $8e^{i(11\pi/6)}=8e^{-i(\pi/6)}$

15. $2e^{i(4\pi/3)}=2e^{-i(2\pi/3)}$

17. -2 **19.** $-\sqrt{2}/4-i(\sqrt{2}/4)$

21. $-2\sqrt{3}+2i$

23. $-\frac{3}{2}-\frac{3}{2}\sqrt{3}i$

25. $\cos 1+i \operatorname{sen} 1 \approx 0.5403+0.8415i$

27. $4-6i$ **29.** $7i$

31. $2e^{-i(\pi/7)}$ **33.** $3e^{i(4\pi/11)}$

35. Si $z = \bar{z}$, entonces $\alpha + i\beta = \alpha - i\beta$, o bien $i\beta = -i\beta$, lo cual es posible si y sólo si $\beta = 0$ de manera que z es real. Si z es real, entonces $z = \alpha = \bar{z}$.

37. $z\bar{z} = (\alpha + i\beta)(\alpha - i\beta) = \alpha^2 - (i^2\beta^2) = \alpha^2 + \beta^2 = |z|^2$

39. El lugar geométrico de los puntos que están sobre una circunferencia en el plano complejo, con centro en z_0 y radio a. Si $z_0 = x_0 + iy_0$, entonces, en las coordenadas x y y, se trata de una circunferencia cuya ecuación es $(x - x_0)^2 + (y - y_0)^2 = a^2$.

41. Supóngase que $p(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1z + a_0 = 0$. Entonces

$\overline{z^n+a_{n-1}z^{n-1}+\cdots+a_1z+a_0} = \bar{0} = 0 = \overline{z^n} + \overline{a_{n-1}z^{n-1}} + \cdots + \overline{a_1z + a_0} = \overline{z^n} + a_{n-1}\overline{z^{n-1}} + \cdots + a_1\bar{z} + a_0$ (puesto que los elementos a_i son reales) $= \bar{z}^n + a_{n-1}\bar{z}^{n-1} + \cdots + a_1\bar{z} + a_0 = p(\bar{z}) = 0$. Aquí hemos empleado el hecho de que para cualquier entero k, $\overline{z^k} = \bar{z}^k$. Esto se deduce fácilmente si expresamos z en forma polar. Si $z = re^{i\theta}$, entonces $z^n = r^ne^{in\theta}$, $\overline{z^n} = r^ne^{-in\theta}$, $\bar{z} = re^{-i\theta}$, y $\bar{z}^n = r^ne^{-in\theta} = \overline{z^n}$.

43. Como $(\cos\theta + i \operatorname{sen}\theta)^1 = \cos 1\cdot\theta + i \operatorname{sen} 1\cdot\theta$, la fórmula de De Moivre se verifica para $n = 1$. Considere que es válida para $n = k$. Esto es, $(\cos\theta + i \operatorname{sen}\theta)^k = \cos k\theta + i \operatorname{sen} k\theta$. Entonces $(\cos\theta + \operatorname{sen}\theta)^{k+1} = (\cos\theta + i \operatorname{sen}\theta)^k(\cos\theta + i \operatorname{sen}\theta) = (\cos k\theta + i \operatorname{sen} k\theta) \times (\cos\theta + i \operatorname{sen}\theta) = [\cos k\theta \cos\theta - \operatorname{sen} k\theta \operatorname{sen}\theta] + i[\operatorname{sen} k\theta \cos\theta + \cos k\theta \operatorname{sen}\theta] = \cos(k\theta + \theta) + i \operatorname{sen}(k\theta + \theta) = \cos(k + 1)\theta + i \operatorname{sen}(k + 1)\theta$, lo cual es la fórmula de De Moivre para $n = k + 1$.

Índice